A Kinematic Model for Constructed Dialog in American Sign Language

John C. McDonald[1], Rosalee Wolfe[1], Robyn Moncrief[1], Souad Baowidan[1], Jerry Schnepp[2],

[1]DePaul University, Chicago IL {jmcdonald,wolfe}@cs.depaul.edu, {rkelley5, sbaowida}@mail.depaul.edu
[2]Bowling Green State University, Bowling Green, OH schnepp@bgsu.edu

Avatars that portray sign language hold great promise for improving Deaf / hearing communication in their application in education and automatic translation. Although avatars can portray legible sign language from phonemic tags alone (Wolfe, Cook, McDonald, & Schnepp, 2011), producing believable, natural motion requires the model to also compute a number of sub-phonetic movements that are not included in linguistic annotation. These include subtleties in individual joint timings which, while not part of the linguistic structure, are essential to achieving lifelike, communicative animations of the discourse. These subtleties come in several forms including the timing of anticipatory joint motions that cascade through the body as muscles contract, and compensatory motions that facilitate the maintenance of balance. In fact, the referential shift for constructed dialog is an important case study for such processes, because it affects the entire body.

Referential shifting is a narrative technique often used in American Sign Language (Metzger, 1995; Padden, 1986). To cue constructed dialog, a signer uses a referential shift to assume the role of a protagonist in the discourse (Lillo-Martin, 1995; Lee, Neidle, MacLaughlin, Bahan, & Kegl, 1997; Quer, 2011). The movement depends on the referents' locations in sign space, and their interaction including dialog (Wilbur & Patchke, 1998). For purposes of automatic translation, this is best synthesized via a kinematic model rather than resorting to prerecorded or captured clips. Any synthesis system must support referential shifts in a way that does not interfere with other co-occurring linguistic processes.

The purpose of such a model is to synthesize the loci and timing of constructed dialog from linguistic tags. These include referents, eye gaze, eyelid aperture and other associated nonmanual signals (Rogers, 2011). The model must accommodate all three representational spaces described in (Morgan, 1999), including narrator space, fixed referential framework and shifted referential framework. In particular, the last of these requires the computation of torso movements necessary to orient the signer to indicate both protagonists in the dialog.

The need for these features is widely recognized in the animation industry and professional animators are highly adept at incorporating them in hand animation. Unfortunately, this is time consuming even for experienced animators. Automating the computation of these features would result in more efficient synthesis. Our new model is general enough to support both sign language animation and gesture that accompanies spoken discourse.

Our model's application to ASL is based on studies of annotated corpora (Neidle, 2001; Poor, 2008), and infers the necessary joint orientations from linguistic tags to create the referential shift movement. It automatically computes staggered timing and compensatory joint motion. The model also cooperates with other co-occurring linguistics processes such as verb agreement, lexical modifiers, and formation of questions. Output from this kinematic model is visualized via a signing avatar. In our presentation we will discuss the details of the model and report the results of our acceptability testing with the Deaf community.

References

Lee, R., Neidle, C., MacLaughlin, D., Bahan, D., & Kegl, J. (1997). Role shift in ASL: A syntactic look at direct speech. In C. Neidle, D. MacLaughlin, & R. Lee (Eds.), *Syntactic Structure and Discourse Function: An Examination of Two Constructions in ASL* (pp. 24-45.). Boston, MA: American Sign Language Linguistics Research Project.

Lillo-Martin, D. (1995). The Point of View predicate in American Sign Language. In I. K. (Eds.), *Language, Gesture, and Space* (pp. 155-170). Hillsdale, NJ: Lawrence Erlbaum Associates.

Metzger, M. (1995). Constructed Dialogue and Constructed Action in American Sign Language. In C. Lucas, *The Sociolinguistics in Deaf Communities* (pp. 255-271). Washington, D.C.: Gaullaudet University Press.

Morgan, G. (1999). Event packaging in British Sign Language discourse. In E. Wiinston (Ed.), *Storytelling & Conversation: Discourse in Deaf Communities* (pp. 27-58). Washington, DC: Gallaudet University Press.

Neidle, C. (2001). SignStream™: A database tool for research on visual-gestural language. (B. Bergman, P. Boyes-Braem, T. Hanke, & E. Pizzuto, Eds.) *Sign Language & Linguistics , 4* (1/2), 203-214.

Padden, C. (1986). Verbs and role-shifting in ASL. *Proceedings of the fourth national symposium on sign language research and teaching*, (pp. 44-57).

Poor, G. (2008). *ASL Video Dictionary and Inflection Guide* . Retrieved May 15, 2011, from http://www.ntid.rit.edu/dig

Quer, J. (2011). Reporting and quoting in signed discourse. In E. Brendel, J. Meibauer, & M. Steinbach (Eds.), *Understanding Quotation* (pp. 277-302). Berlin: Mouton de Gruyter.

Rogers, K. L. (2011). *American Sign Language Verb Categories in Constructed Action (Doctoral Dissertation).* Retrieved from ProQuest Dissertation and Theses (Accession Order No. AAT 3516286).

Wilbur, R., & Patchke, C. (1998). Body leans and the marking of contrast in American Sign Language. *Journal of Pragmatics , 30*, 275-303.

Wolfe, R., Cook, P., McDonald, J., & Schnepp, J. (2011). Linguistics as structure in computer animation: Toward a more effective synthesis of brow motion in American Sign Language. (A. Hermann, & M. Steinbach, Eds.) *Nonmanuals in Sign Language Special issue of Sign Language & Linguistics , 14* (1), 179-199.