A computational model of role shift to support the synthesis of signed language (English)

John McDonald, Rosalee Wolfe, Robyn Moncrief, Souad Baowidan,
School of Computing, DePaul University, Chicago, IL
{jmcdonald, wolfe}@cs.depaul.edu, {rmoncrie, sbaowida}@mail.depaul.edu

This work addresses the theoretical issue of computational modeling for signed language avatars. Computer-generated avatars that portray signed language have the potential to improve Deaf/hearing communication through their application to Deaf education, interpreter training and automatic translation. However, current signed language generation systems are only capable of producing simple utterances because they lack complete, accurate computational models of syntactic processes. A case in point is the theoretical issue of developing a generalized computational model for role shift.

Role shifts in signed languages convey constructed dialog that reports on the statements, actions or thoughts of someone other than the person who is signing (Lillo-Martin, 2012). A computational model of role shift needs to incorporate reported dialog (Friedman, 1975), constructed action (Liddell & Metzger, 1998), and a framework of narrator, fixed referential and shifted referential spaces (Morgan, 1999). To be complete, such a model would also need to accommodate differences in scale among the spaces (Dudis, 2004), (Perniss, 2007) as well as rotation (Janzen, 2004).

Our methodology uses a two stage process. The first stage analyzes annotated corpora to produce a computational model. The second stage evaluates the model by generating an animated story, which is rated by members of the Deaf community. We use the feedback gathered in the second stage to modify the model prior to a second round of evaluation.

Our initial computational model utilized dual mathematical coordinate systems, one in world coordinates and one in local, to accommodate the framework of the three spaces described by Morgan. It also incorporated the perspective shift described by Janzen. It generated a signer's eye gaze, head direction and spinal orientation that are visually necessary to indicate a role shift. This approach improves on previous efforts that used a single coordinate system (Braffort & Dalle, 2008), and can facilitate the synthesis of conversations among multiple protagonists involving a shared referent.

We evaluated the model by generating an animation of a story involving three characters. Twenty-two members of the Deaf community rated the story's sentences for clarity, grammaticality, acceptability and naturalness. The sentences received high ratings for grammaticality and clarity, but the scores for acceptability and naturalness were significantly lower.

We refined the model to incorporate the biomechanics of transitioning from one role to another and regenerated the story. Thirty-six members of the Deaf community evaluated the revised story, see Figure 1. The most important difference was in naturalness, which shifted positively as confirmed by a Mann-Whitney rank test ($p < .05$). Acceptability and grammaticality showed no significant difference, and while clarity was slightly lower, the results were still positive. The evaluation of the revised computational model also showed increased stability with all positive interquartile ranges.

In conclusion, we developed a computational model characterizing a role shift capable of producing synthesized utterances of signed language that were judged as clear, grammatical and acceptable by members of the Deaf community. The model infers transitions from linguistic processes and requires no additional specification for sentence generation.

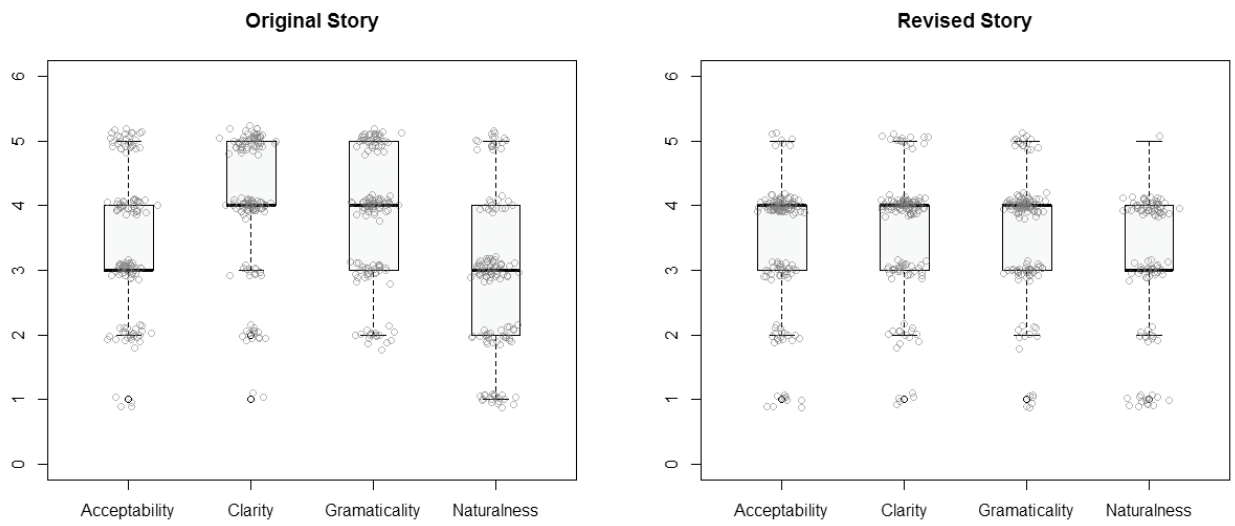

Figure 1: Results of the user tests for the original and revised versions of the story on a 5-point Likert scale.

Bibliography

Braffort, A., & Dalle, P. (2008). Sign language applications: preliminary modeling. *Universal access in the information society, 6*(4), 393-404.

Dudis, P. (2004). Body partitioning and real-space blends. *Cognitive Linguistics, 15*(2), 223-238.

Friedman, L. (1975). Space, Time and Person Reference in American Sign Language. *Language, 51*(4), 940-961.

Janzen, T. (2004). Space rotation, perspective shift and verb morphology in ASL. *Cognitive Linguistics, 15*(2), 149-174.

Liddell, S., & Metzger, M. (1998). Gesture in sign language discourse. *Journal of Pragmatics, 30*, 657-697.

Lillo-Martin, D. (2012). Utterance reports and constructed action. In R. Pfau, M. Steinbach, & B. Woll (Eds.), *Sign Language: An International Handbook HSK 37* (pp. 365-387).

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.

Perniss, P. (2007). Achieving spatial coherence in German Sign Language narratives: The use of classifiers and perspective. *Lingua, 117*(7), 1315-1338.