



HAL
open science

Corpus Design for Signing Avatars

Kyle Duarte, Sylvie Gibet

► **To cite this version:**

Kyle Duarte, Sylvie Gibet. Corpus Design for Signing Avatars. Workshop on Representation and Processing of Sign Languages: Corpora and Sign Language Technologies, May 2010, Valetta, Malta. pp.1-3. hal-00505182

HAL Id: hal-00505182

<https://hal.science/hal-00505182>

Submitted on 22 Jul 2010

HAL is a multi-disciplinary open access archive for the deposit and dissemination of scientific research documents, whether they are published or not. The documents may come from teaching and research institutions in France or abroad, or from public or private research centers.

L'archive ouverte pluridisciplinaire **HAL**, est destinée au dépôt et à la diffusion de documents scientifiques de niveau recherche, publiés ou non, émanant des établissements d'enseignement et de recherche français ou étrangers, des laboratoires publics ou privés.

Corpus Design for Signing Avatars

Kyle Duarte, Sylvie Gibet

Université de Bretagne-Sud, Laboratoire VALORIA, Vannes, France
kyle.duarte@univ-ubs.fr, sylvie.gibet@univ-ubs.fr

Abstract

The SignCom project uses motion capture (mocap) data to animate a virtual French Sign Language (LSF) signer. An important part of any signing avatar project is to ensure that a computer animation engine has a large quantity of interesting and on-topic signs from which to build novel signing sequences. In this article, we detail the process of selecting an adequate range of signs and situations to be included in our corpus: from controlling discourse topic to including signs that can accept modified movements or handshapes, we describe how an avatar corpus has a different motivation than traditional signed language corpora.

1. Introduction

Though the field of signed language corpus building is young, designing a corpus specifically for application in signing avatars already requires a deviation from available standard practices. Often, corpora for sign retrieval are based on semi-scripted interactions that yield many instances of a restricted set of signs, useful for building unique dialogs later on. We describe here the considerations we have taken in designing the SignCom signing avatar corpus and how they might vary from corpora designed solely for linguistic analysis.

2. Previous Research

Sinclair defines a computer corpus as “a corpus which is encoded in a standardised and homogenous way for open-ended retrieval tasks.” These structures are evident in the Australian, British, Dutch, Greek, and other signed language corpora: language samples are coded (usually with ELAN) to indicate phonology, morphosyntax, and other language components for later retrieval and analysis (Johnston and de Beuzeville, 2009; Crasborn and Zwitserlood, 2008; Efthimiou and Fotinea, 2007).

However, where these corpora have been developed to serve as living representations of signed languages across individuals and time, corpora used in the realm of language synthesis attempt to find a restricted sample of language that can be reused in a variety of scenarios.

Akin to digital motion databases that might, for example, index the movements of a basketball player, signing avatar corpora require many repetitions of the same sign in different contexts to provide an interesting base for research and animation. With multiple phonological instances of the same sign recorded, a computer animator can choose a best-fit sign out of many, instead of forcing a single instance of the sign into a novel context. These principles have shaped the range of French Sign Language (LSF) signs made available in the SignCom project.

3. SignCom Corpus Design

The SignCom corpus has been designed by a team of researchers that includes linguists and computer scientists, hearing and Deaf. With multiple points of view converging on solving a multidisciplinary problem, several opposing goals have had to be weighed for our desired outcome.

Three excerpts from the segments we have used most often for language synthesis to date are shown below; after, follow descriptions of our opposing goals and our eventual solutions.

Last Saturday evening I organized a cocktail party. I invited some friends over to my house. In order to facilitate communication, I pushed the chairs in the living room into a semi-circle. There was a coffee table for our drinks, and an American bar with various drinks, fruits, glasses, and straws.

I asked my friend, “what do you want?”
(S)he said, “I would like vodka and orange juice.”
“Okay,” I responded. I selected a tall thin glass and added vodka about a quarter of the way up. I filled the rest of the glass with orange juice and handed it to my friend.

I asked the next friend what (s)he wanted.
(S)he responded, “eh, I like any drink, so I don’t really know. What do you suggest?”
“I’d suggest a cocktail named *Cuba Libre*,” I said.
“What’s inside that?” (s)he asked.

I said it would be a surprise. I got a tall glass and added a couple of ice cubes. I poured a little lemon juice in the glass, added some rum to that, and filled the glass with cola, then served it to my friend.

All in all, I was quite happy that the evening went well.

3.1. Depth vs. Breadth and Variation vs. Consistency

Traditional corpora attempt to gather a large number of signs to represent the largest slice possible of a language. For the purposes of language synthesis, however, the researcher wants to have control over the types that appear in the corpus, and would prefer several tokens of these types. Dialogues are thus preplanned to ensure multiple instances of a single type are available for searching and retrieval in later experiments, also allowing for best match selection among token candidates.

The SignCom corpus contains three thematic sections: the Cocktail story, and the Galette and Salad interactions.

These themes limit the material that can be discussed in elicitation sessions to a narrow vocabulary. Discussed in long interactions, the signer provides a large number of tokens relative to the narrow focus. The Cocktail story section, measuring roughly one third of the overall corpus, contains the tokens shown in Table 1, among others. With this variety and frequency of cocktail-related lexemes, we are able to produce a number of novel utterances around the same subject. For example, Figure 1 shows a sequence we have constructed from various single signs and sign phrases. The final result is interpreted as

I asked the next friend what (s)he wanted.
 (S)he responded, “eh, I don’t like fruity drinks, so I don’t really know. What do you suggest?”
 “I’d suggest a cocktail named *Cuba Libre*,” I said.
 I gave it to her and (s)he took it.
 “Great!”

Note that constructing this utterance requires selecting signs from various parts of the corpus. The movements of two signs were inverted phonologically to evoke a contrary meaning. The purposeful inclusion of such directional signs was intended for such an utterance, and is detailed in Section 3.3., below.

Finally, as there is a necessary balance of control within variability for avatar projects, signing avatar corpora do not provide the level of variation needed for a sociolinguistic study.

Table 1: The tokens of highest occurrence in the Cocktail story section of the SignCom corpus.

14x WHAT	7x WANT
9x VARIOUS	4x FILL
8x COCKTAIL	3x JUICE
8x DRINK (n.)	3x ORANGE
8x EVENING	3x VODKA
8x FRUIT	2x RUM
8x POUR	2x SUGGEST
7x GLASS	

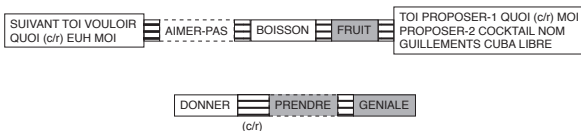


Figure 1: Signs can be rearranged to create novel phrases. Here, signs are retrieved from two different recording takes (white and gray backgrounds) and linked with transitions created by our animation engine (striped background). The sign AIMER (“like”) is reversed to create AIMER-PAS (“dislike”), as is DONNER (“give”) to create PRENDRE (“take”). Finally, a role shift, shown as (c/r), is included in one transition to ensure discourse accuracy and comprehension.

3.2. Open-Ended vs. Scripted

Anonymity in contributions to signed language corpora has been an important conversation within the Deaf communities that support this type of research. At the most basic level, given the face’s active involvement in the signing event it is impossible to hide the identity of the signer. Linguistic data has thus been subject to tight controls regarding rights releases to allow data analysis among researchers, as well as data publishing to wider and/or public audiences. This topic becomes even more sensitive when open-ended questions are used to elicit stories for linguistic corpora. Existing corpora use guiding topics to elicit personal responses, which may include reports of abuse or other illegal activities; eventually such data would require censorship when making corpora public. As signing avatars almost inevitably become publicly viewable, researchers aim to avoid controversial topics in recording sessions. As an added benefit, the avatar medium aides in anonymizing elicited data by providing a new face and body for the signer. Figure 2 shows our language consultant alongside the avatar that replays her signing in our animation system.

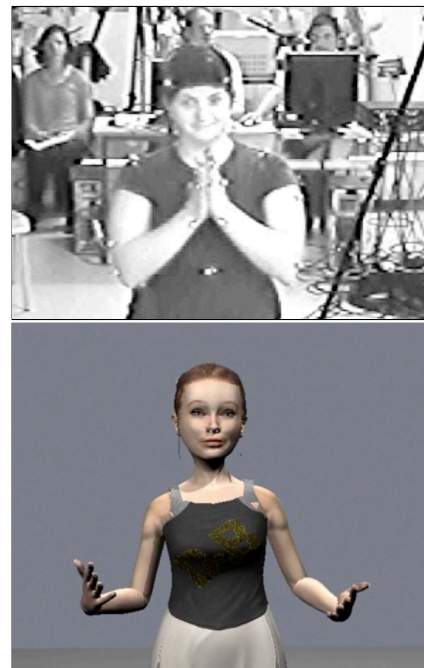


Figure 2: Avatars provide new identities to signers without covering the face, an important articulator for the signing event.

3.3. Experiments in altering phonological components

Our specific research interests brought us to include a number of indicating verbs and depicting verbs in the SignCom corpus. Among our scientific inquiries are the questions of whether playing reversible indicating verb motions backwards will be convincing and whether altering the hand-shape of a stored depicting verb will be understood as a change in meaning.

For example, an LSF signer can reverse the movement of the LSF sign AIMER (“like”) to produce the meaning

AIMER-PAS (“dislike”), however human motion theories predict that the motion profile of the reversed sign AIMER-PAS will not be a simple inversion of the forward sign AIMER. By creating sequences that include the captured sign AIMER played backward in a computer-generated utterance, we can perform simple perception tests with signers to judge the acceptability of this relatively straightforward animation technique. We believe these inversions will be understood by signers, so we have included them in our corpus to challenge existing understandings of human motion.

In addition, given our inclusion of signs that take multiple handshapes, like DONNER (“give”), we can substitute handshapes from other signs to influence meaning. In the case of DONNER, most often sign in our corpus as if the signer is handing a glass to someone, a handshape substitution could yield additional meanings, such as giving a piece of paper or giving something dirty (paired with an appropriate facial expression).

3.4. Technical Considerations

Finally, avatar systems must incorporate motion capture (mocap) files that represent the movement of the body, generally much more compact than video files. This incorporation, as well as results of our avatar corpus, is detailed in the paper “Heterogeneous Data Sources for Signed Language Analysis and Synthesis” presented at the LREC 2010 main conference (Duarte and Gibet, 2010).

4. Conclusion

In all, creating databases of signs for signing avatars is not unlike some aspects of traditional linguistic corpora. However, key factors such as dialogue content and style, as well as technical inclusions, must be considered in designing an avatar corpus.

For the SignCom project, we have centered our elicitation sessions around three themes so as to limit the scope of vocabulary attained, and increase the tokens available to us for creating similarly-themed novel utterances. By studying semi-scripted stories, we virtually eliminate the possibility that the signer provides sensitive information that should be held from the public’s view, and better control the corpus’s content for later retrieval. By the nature of animating an avatar, we preserve anonymity for our signer.

Other project goals brought us to include a number of signs that could exist with altered movements or handshapes, to test our animation system’s ability to interchange body parts across signs, as well as to better understand signers’ perception and comprehension of signing avatars.

Having collected our data, we believe that we have an excellent base with which we can create convincing animations of French Sign Language, due in large part to the intentional way we built the SignCom corpus.

5. Acknowledgements

We would like to thank Julia Pelhate (Websourd, Toulouse) and Juliette Dalle for their participation in our signed language mocap recordings, as well as Patrice Dalle (IRIT, Toulouse) for his support on the corpus dialogue design.

This project is funded by the *French National Research Agency*.

6. References

- Onno Crasborn and Inge Zwitterlood. 2008. Annotation of the video data in the Corpus NGT. Technical report, Department of Linguistics and Center for Language Studies, Radboud University, Nijmegen, the Netherlands, November.
- Kyle Duarte and Sylvie Gibet. 2010. Heterogeneous data sources for signed language analysis and synthesis. In *Proceedings of the 7th International Conference on Language Resources and Evaluation, Valletta, Malta*, 19-21 May. To appear.
- Eleni Efthimiou and Stavroula-Evita Fotinea. 2007. GSLC: Creation and annotation of a Greek Sign Language corpus for HCI. In *Universal Access in Human Computer Interaction. Coping with Diversity*, volume 4554 of *Lecture Notes in Computer Science*, pages 657–666. Springer, Berlin.
- Trevor Johnston and Louise de Beuzeville. 2009. Researching the linguistic use of space in Auslan: Guidelines for annotators using the Auslan corpus. Technical report, Department of Linguistics, Macquarie University, Sydney, June.