



**HAL**  
open science

## Transferability of new affordances in collaborative analyses of complex data?

Kristine Lund, Elizaveta Chernyshova, Laurène Smykowski, Lydia Heiden, Thomas Franco Pinto, Heike Baldauf-Quilliatre, Isabel Colón de Carvajal, Carole Etienne

### ► To cite this version:

Kristine Lund, Elizaveta Chernyshova, Laurène Smykowski, Lydia Heiden, Thomas Franco Pinto, et al.. Transferability of new affordances in collaborative analyses of complex data?. Science of Team Science 2021, International Network for the Science of Team Science and Virginia Tech, Jun 2021, Blacksburg, Virginia (on-line), United States. hal-04035758

**HAL Id: hal-04035758**

**<https://hal.science/hal-04035758v1>**

Submitted on 27 Mar 2023

**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.



SciTS

Building the knowledge base  
for effective team science

## Transferability of new affordances in collaborative analyses of complex data?

Kristine Lund<sup>123</sup>, Elizaveta Chernyshova<sup>23</sup>, Laurène Smykowski<sup>123</sup>, Lydia Heiden<sup>23</sup>, Thomas Franco Pinto<sup>23</sup>, Heike Baldauf-Quilliatre<sup>23</sup>, Isabel Colon-Carvajal<sup>23</sup>, Carole Etienne<sup>23</sup>  
CNRS<sup>1</sup>, Ecole Normale Supérieure de Lyon<sup>2</sup>, Université Lumière Lyon 2<sup>3</sup>  
Kristine.lund@ens-lyon.fr

Hello to all. I'm Kristine Lund and we are from Lyon, France. Our work on collaborative analysis of complex data is carried out in the context of a 5-year project that I lead called ASLAN, involving around 200 researchers and staff, with a 4.3 million-euro budget.

ASLAN studies the complexity of language. One of our projects is DatAgora, and its goal is to work with complex, multimodal, multiactivity traces of human interaction. We are at the crossroads of language sciences, education, and computer science. Our work is transdisciplinary because we involve stakeholders outside of research. We collaborate with teachers and students, with migrants and language users around the world, with patients and health professionals, with families and associations, with policy makers, both public and private sectors, and so with citizens, society and government.

In all our projects, there must be an epistemological positioning at the outset in that theory informs how data is collected, how it's prepared, visualized, analyzed, and interpreted. Often this loop repeats in that results may lead to new data collection after researchers revisit the way they view their research.

Our overarching objective is to scaffold activities 3,4,5 & 6 in this loop in order to facilitate this revisiting. To this end, we are exploring the use of collaborative software SAGE2 on a wall of screens.

You can see our set-up on the left. We have 4 screens connected to a PC running windows and SAGE2, with a sound bar, webcam, and multidirectional microphone. Multiple people can connect to the space, also from a distance. And multiple objects can be projected dynamically.

Our more specific goals are three-fold, and they occur with these nested projects. The initial team should explain to others the basic usage of SAGE2 and the wall of screens. And they should test usability and our scaffolding objective. They did this by brainstorming the preparation for a conference they submitted to, within the context of the project ENJEUX. So, you can see how these projects are nested. Enjeux is a project in interactional linguistics, DatAgora is a Team Science Project on Data Visualization, and ASLAN is the overarching project on Language Complexity. We wanted to keep in mind the extent to which what we found in the ENJEUX project could be applied to other projects within ASLAN that could also be included within DatAgora.



SciTS

Building the knowledge base  
for effective team science

Our approach is a Snowball method. It is what it sounds like. We want to bring new users from ASLAN on board through word of mouth and initial success. The first group - Lisa, Laurène & colleagues tested SAGE2 and the wall of screens for their project ENJEUX in interactional linguistics, they reported back and shared their best practices to both the ENJEUX project and the DatAgora project. They implemented the feedback they got into how they continued their tests and the cycle started again. A glossary of terms was necessary — here in French — in order to understand how SAGE2 and the wall of screens can be piloted by a laptop.

In interactional linguistics, complex data means synchronized data coming from different sources and representing different simultaneous activities. So, we have a video, top left, and software in which the video is also embedded, that shows a synchronized and coded transcription of what is said and the gestures that are made (top right). Then on the bottom right we have a visualization in histogram form of analyzed data concerning how gaze is aligned. And on the bottom left is a simple post-it note, coupled with a drawing tool that allows researchers to make links between the different data representations. The orange arrows show how elements are related. For example, the circle in the top right of the temporal, coded representation of the video and transcription corresponds to the histogram representation of the type of gaze alignment on the bottom right.

So now that you know a little bit of what an interactional linguistic pays attention to, we can move to how they mobilized the wall of screens to do their data analysis. Initially, they organized the space. On the left they decided to implement a dynamic workspace where they could move data in and out. On the right they put visualizations that were transversal to the videos and transcriptions thereof that they were reviewing, such as a timeline describing phases of the card playing, the interaction they were analyzing.

In the dynamic space they miniaturized pairs of video extracts and their associated transcriptions (upper left). They also added gifs that had been pre-prepared that showed specific types of movement. Here you see how they opened video-transcription pairs one by one in order to review them, think about how to categorize them, re-miniaturize them, and review the next one.

This view is a distance view where we are following the presentation in zoom. On the bottom there is a temporal visualization where researchers can keep track of where in the sequence the video extract occurred, in order to maintain understanding of the global game context.

And here - in a re-organization of the space according to their needs, two post-it notes served as an area in which to categorize the video-transcription pairs. Four were extracts in which players referenced physical objects in their talk and gesture (in the blue post-it, above) and four were extracts in which players referenced abstract concepts such as rules (in the peach colored post-it, below).



SciTS

Building the knowledge base  
for effective team science

in summary then, what did our explorers in interactional linguistics do and how can it be useful for other projects that work on complex data of a different sort in ASLAN?

- They mobilized affordances of the workspaces in relation to their analytical needs
- Their complex data was organizable in a way that helped their process
- We already know from past studies that simultaneous manipulation led to spontaneous or orchestrated role distribution
- and finally, the set up facilitated trial-and-error data categorization

These are sufficiently general to encourage us to pursue our work in different analytical contexts. We are also inspired by the Decision Theatre in Arizona, where sufficient data has led to visualization models that are applicable to multiple contexts. But we are only at the very beginning. Negative points include much work needing to be done beforehand to be operational.

We would like to argue that for Science of Team Science, it is useful to collate and analyze how we can characterize visualization, analysis and interpretation of complex data across different project contexts. We'll document our snowball method and describe how the scaffolding we propose can contribute to a re-positioning of epistemological approach at the crossroads of language science, education, and computer science. So, please stay tuned, as we'll also be thinking about similar work in other disciplines!