

## SHORT TERM SCIENTIFIC MISSION (STSM) SCIENTIFIC REPORT

This report is submitted for approval by the STSM applicant to the STSM coordinator

**Action number: CA19102**

**STSM title: Literacy and speech comprehension in the era of mediatized language**

**STSM start and end date: 13-06-2021 to 12-07-2021**

**Grantee name: Chiara Celata**

### PURPOSE OF THE STSM:

(248 words)

The purpose of my STSM project was to develop collaborative and interdisciplinary knowledge on the impact of new technologies on linguistic behavior in different groups of speakers. More specifically, my focus was on the literacy variable, and on the need of (1) developing innovative ways of measuring literacy in the age of mediatized cultures, (2) understanding the impact of how the emergence of ‘internet varieties’ impacts on how language is perceived and produced, with a particular emphasis on speech (phonetic variation and phonological change).

Given these research questions, I applied for a STSM to the MoDyCo lab for the following reasons.

First, I was primarily interested in behavioral and neurophysiological methods to track language comprehension and production in speakers with different literacy levels. MoDyCo provides an excellent base of knowledge and experience, including access to relevant technologies, in the domain of the behavioral and neuro-imaging study of executive control in bilinguals and sensitive populations such as children with different language deficits, and with an expertise also in the domain of multimodal processing of language (see details in my Application).

Second, MoDyCo features an established tradition in corpus-based and data-driven language studies, allowing a solid descriptive base which takes into account the reality of language uses, whether conversational, discursive or textual. This dimension would be important to acquire a critical knowledge of traditional sociolinguistic methodologies in corpora construction and analysis, which will allow to develop more ecological ways of treating the literacy variable as a complex variable.

### DESCRIPTION OF WORK CARRIED OUT DURING THE STSMS

(842 words)

1) Preparatory work.

Virtual meetings with the lab members began well before the time of the STSM. We identified a well-defined set of research questions to tackle the problem of language processing in speakers with different literacy levels, and we decided to focus at the beginning on language comprehension. Experimental paradigms in the behavioral and neurophysiological domains allowing investigation of language comprehension from an oral + multimodal oral-visual perspective were discussed. As a consequence of such preparatory work, we decided to implement a pilot study with French as the target language. More specifically, we decided to analyse the processing of categorical and variable liaison by looking at brain electrical activity through EEG during a listening task, potentially coupled (in the case of a subset of the

participants) to a simultaneous silent reading task.

2) Working out the purposes and methodologies of the pilot study.

The subject of the investigation was established in French liaison. French liaison is a sandhi phenomenon that obeys an extraordinarily high number of linguistic and communicative parameters, and that is high in the meta-linguistic awareness of the French speakers since it generates an evident mismatch between phonology and orthography. One of these, which is however still poorly investigated, is the relationship that the speakers establish with the written form of words while hearing (or uttering) the speech flow containing liaison. Moreover, we know a lot about constraints and variation in liaison production, but almost nothing about comprehension and meta-linguistic judgments of the speakers about variable uses of liaison. Therefore, the first part of the STSM was devoted to establish the purposes and methodologies to be used for the pilot study on literacy and liaison comprehension. We defined the following purposes: (i) to evaluate the speakers' reactions to *violations* of liaison patterns that are in the signal; (ii) to develop an experimental setting allowing to test whether these reactions change when a multimodal oral + written input is provided. We defined the following methodology: a comprehension test with non-invasive recording of brain electrical activity through EEG.

3) Setting out the pilot study: materials, experimental setting, first trials.

We conceived of an innovative experimental setting in which French participants listened to a discourse uttered in their native language during a dozen of minutes. The discourse had been previously uttered and recorded by a professional native French reader. In such text, a series of liaison contexts was included. Each context occurred twice in the text: the first occurrence obeyed the general liaison pattern, the second contained a violation. The participants were asked to simply listen to the text (and reply to some short questions at the end, to ensure that their level of attention was constant throughout the listening). The participants wore the relevant EEG equipment for the measurement of their electrical activity during the task. This design is innovative in proposing an ecological setting, according to which the participants are not exposed to isolated word-like stimuli or sentences, as in traditional experiments, but rather to a continuous text, that is similar to a journalistic essay and therefore easily experienced and processed by the participants. Moreover, in order to evaluate brain reactions in a multi-modal input condition, we envisaged the possibility of adding a silent reading task, e.g. by asking the participants to read the text flowing on a computer screen simultaneously to the oral input. In order to avoid artifacts in the brain signal due to head or face movements during reading, we started investigating different ways of presenting the written input and evaluating pros and cons of each of them. The pilot study was put in place during the STSM and a few trials were made.

4) Discussions with other lab members, planning of future studies and collaborations.

In parallel to the experimental work, the STSM was also devoted to discussion with various lab members about innovative ways of measuring literacy in different populations of speakers. In this respect, particularly fruitful were the meetings with (i) PhD students, scientists and clinicians working on dyslexia and other verbal deficits; (ii) linguists and technical staff working on the "Phonologie du français contemporain" - PFC corpus. With respect to the first point, I could make a precise review of all formal and informal tests that are currently used in the clinical setting to evaluate not just the literacy level of the speakers, but also correlated variables such as reading experience, familiarity with social media and others. This was of fundamental importance to start establishing new ways of quantification of the "reading experience" that might be applied also to non clinical populations (such as foreign language learners, migrants, people from disadvantaged social and economic backgrounds). With respect to the second point, I developed a more detailed knowledge of the PFC corpus and of the sociolinguistic parameters that are taken into account in the surveys which constitute the speech database. This information is crucial in particular to establish more fine-grained categories of liaison patterns in contemporary French, which will be used to raise more specific working hypotheses in the experimental study.

#### **DESCRIPTION OF THE MAIN RESULTS OBTAINED**

(162 words)

As already mentioned in the Description of the work carried out during the visit, there were several outcomes from this STSM. A short summary is provided below:

- in order to set up an experimental approach to the issue of literacy and speech comprehension, we defined a specific experimental setting (i.e. continuous listening with concomitant brain reaction recordings) that is

innovative in some respects and that can be further modeled and enriched to take into account the presence of multi-modal input (with the simultaneous reading of flowing written input);

- we defined the linguistic variables and prepared the experimental materials (stimuli) to be used in the pilot study;
- we did some trial recordings;
- we discussed with several lab members about literacy from a sociolinguistic and clinical perspective, as well as about corpus studies on French liaison and sociolinguistic variation in contemporary French, and this discussion directly fed the implementation of the current pilot study and the definition of future studies/collaborations.

#### **FUTURE COLLABORATIONS (if applicable)**

(286 words)

This STSM is the first of planned exchanges with MoDyCo members. The pilot study that was set up on the occasion of this STSM will be implemented starting from September 2021, which will require additional meetings and strict collaboration during several months. The expected outcomes are then two. First of all, at least one joint publication and one conference presentation. Second, the development of a larger research project, including other researchers from Italy, France and other European countries, addressing the general theme of the role of language experience in speech and language processing. We think that it is urgent to understand how different levels and modalities of experience with language impact on the way the speakers deal with language complexity, language variation and their own communicative/societal needs. Increasing our knowledge on this subject will have a strong impact primarily at the educational level, particularly in multilingual societies, and then at the level of all policies in which human and societal differences are to be managed. Both language complexity and language variation have exponentially increased after the introduction of language technologies and mediatized cultures, and this is why it is essential to frame this kind of investigation within a more general and comprehensive framework of scientific contents and challenges, such as that provided by LITHME. We also think that behavioral and neurophysiological studies are a primary source of information in this respect, provided that they are coupled with detailed accounts of sociolinguistic (and possibly sociological) variables tapping on the linguistic behavior of the speakers. This evidently requires high levels of inter- and multidisciplinary in research projects, which cannot be achieved without the support of large-scale collaborative initiatives putting researchers from different backgrounds in contact with each other.