

# **LEarning and TEaching corpora (LETEC): data-sharing and repository for research on multimodal interactions**

Ciara R. Wigham, Thierry Chanier

► **To cite this version:**

Ciara R. Wigham, Thierry Chanier. LEarning and TEaching corpora (LETEC): data-sharing and repository for research on multimodal interactions. WorldCALL 2013, Jul 2013, Glasgow, United Kingdom. edutice-00778274v1

**HAL Id: edutice-00778274**

**<https://edutice.archives-ouvertes.fr/edutice-00778274v1>**

Submitted on 19 Jan 2013 (v1), last revised 12 Jul 2013 (v3)

**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.

## **LEarning and TEaching corpora (LETEC): data-sharing and repository for research on multimodal interactions**

Ciara R. Wigham & Thierry Chanier

**Keywords:** learning and teaching corpus; data sharing; research data repository; online multimodal interactions; synthetic worlds

The number of online environments language teachers can employ is constantly growing, offering increased potential for L2 interaction analysis. However, research cannot necessarily keep up with technology innovation. One danger is that CALL research will reinvent the wheel each time a new technology emerges. To better understand L2 interaction across different environments, Reffay, Betbeder & Chanier (2012) underline the need to share research situations in formats that allow comparisons between interactions in different online environments to be made and that are open-access.

In the language-learning domain, learner corpora (Granger, 2004; Meunier *et al.*, 2012) are exploited for SLA research. Frequently comprising data from test situations (Reffay *et al.*, 2008) and used in learner-native speaker comparative studies (Botlon, Carter-Thomas & Rowley-Jolivet, 2012), learner corpora focus on learners' productions and consider neither other course participants (tutors, native speakers) nor the learning context.

A LEarning and TEaching Corpus (LETEC) links, following international standards, all elements resulting from an online learning situation (Chanier & Ciekanski, 2010). It comprises a XML "manifest" which describes the corpus' components: the learning design, the research protocol (questionnaires, interview data), the interaction data (audio, textchat, video), all participants' productions and licences relating to ethics and access rights. The XML schema allows interactions from different tools and environments (conceptual map editor, blogs, synthetic worlds...) to be stored and described in a standardized way, facilitating data analysis. Interaction data is included in environment-independent formalisms.

Our presentation will introduce the methodology for building a LETEC. Using data from a CLIL (*Content and Language Integrated Learning*) course which employed the synthetic world *Second Life*, and analyses into nonverbal and L2 verbal interaction in this environment (Wigham & Chanier, in press; Wigham & Chanier, 2012), we will illustrate how LETEC methodology may help sustain CALL research beyond the hype of the latest online environment.