40 DOTI: Databank of Oral Teletandem Interactions

Solange Aranha[1] and Paola Leone[2]

Abstract

This contribution aims at (1) discussing the characteristics of collecting, filing and storing data to have a databank of oral interactions between university students whose main objective is the learning of a second language through teletandem; and (2) defining the steps for further collections and storage. Our data are Skype sessions of foreign language learners who interact via Voice Over Internet Protocol (VOIP) with a proficient partner in the language they are learning. Our databank aims at (1) giving value to teletandem as a situated learning context, (2) substantiating the research carried out in the field, and (3) offering other researchers the possibility to access data to confirm or refute published research. We first define a schema for interpreting teletandem sessions according to the Interaction Space (IS) Model as defined by Chanier and colleagues (2014). Subsequently, we discuss metadata concerning contexts (e.g. description of the university and of the language courses) and learning scenarios (e.g. objectives, materials).

Keywords: teletandem, databank, oral communication, language learning, interaction space model, computer mediated communication.

1. Universidade Estadual Paulista - São José Do Rio Preto, Brasil FAPESP # 2015/02048-6; solangea@ibilce.unesp.br

2. Università del Salento, Lecce, Italy; paola.leone@unisalento.it

How to cite this chapter: Aranha, S., & Leone, P. (2016). DOTI: Databank of Oral Teletandem Interactions. In S. Jager, M. Kurek & B. O'Rourke (Eds), *New directions in telecollaborative research and practice: selected papers from the second conference on telecollaboration in higher education* (pp. 327-332). Research-publishing.net. https://doi.org/10.14705/rpnet.2016.telecollab2016.525

1. Introduction

Teletandem (Vassallo & Telles, 2006) is a form of computer mediated interaction by which two students, proficient in two different languages, interact via VoIP technology and/or via text chat. This telecollaborative practice respects the principles proposed by Brammerts (1996): autonomy, separation of languages and reciprocity. Teletandem is nowadays a teaching/learning context which has been institutionalized in different universities around the world and has become a relevant research field in applied linguistics. Over the years, researchers have been collecting, transcribing and analyzing data in different ways according to the needs of their studies (c.f. www.teletandembrasil.org).

As part of a shared project between UNESP and University of Salento, we are now aiming at building a databank with common characteristics (same methodology of collection and transcription) which may be useful for researchers in planning their tasks within telecollaboration activities, in understanding how telecollaboration works and may be optimized, and in developing linguistic research within telecollaboration environments, among others. Our first step is to apply to teletandem data the IS model (Chanier et al., 2014), by which some researchers are trying to characterize different Computer-Mediated Communication (CMC) genres (mostly written, such as Facebook). IS is defined as "an abstract concept, located in **time […]** where interactions between **a set of participants** occur within an **online location**" (Chanier et al., 2014, p. 5).

Considering that teletandem is organized around various tasks in which a language instructor and a class group are involved, the concept of Learning Scenario (LS) becomes relevant, since it describes different task sequences (Mangenot, 2008; Foucher, 2010). LS helps us determine the characteristics that underlie teletandem practice. In this paper, we show how these concepts (IS and LS) are applied to our data and how they can contribute to define Data of Oral Teletandem Interactions (DOTI) metadata which are mostly created for interrogating the databank.

2. Methodology

At UNESP and at University of Salento, teletandem is not a stand alone practice but it comes together with other tasks, carried out both via Information and Communication Technologies (ICT) and in the classroom. Each teletandem session takes about one hour and occurs once a week. At UNESP, Brazilian students, whose mother tongue is Portuguese, interact with American students, proficient in English. At UNISALENTO, Italian students interact with British students.

Both contexts – UNESP and Unisalento (and partner institutions) – have students from different courses who are learning the language and practising it via teletandem sessions. The levels of proficiency vary and are not a key factor to be enrolled in the activity. Each partnership usually lasts from 8 to 15 sessions, depending on the learning scenario. All participants signed a consent form – developed within the exigencies of each university – for video recording oral sessions[3] which are stored[4].

DOTI contains data from 2012 to 2015, in a total of over 650 hours of conversation (Portuguese and English – Italian and English). Some data have been transcribed. Among other communicative data so far described during conferences and in literature following the IS model, DOTI is peculiar since it is compiled by synchronous multimodal interactions during which different modes are employed for communication (text, gestures, oral, images, etc.). Thus, DOTI data represent a complex environment.

Teletandem interactions are part of different learning scenarios which, in both institutions, are shaped in macro and microtasks (objectives and description). UNESP and Unisalento share the macrotasks' aim which is preparing students to participate actively in (computer mediated) oral interactions with a proficient speaker and be aware of all the linguistic and cultural strategies that such a

3. So far we have been using Evaer, a capture Skype video and audio data to record (see www.evaer.com).

4. In Brazil, a detailed description of storage process can be found in Aranha, Luvizari-Murad and Moreno (2015).

practice involves. In the Brazilian and Italian universities, such an objective is reached via different microtask sequences carried out during mediation sessions and computer mediated oral sessions.

These mentioned features are useful guidelines for defining metadata.

3. Discussion

Some metadata will be presented: first of all, those concerning teletandem as IS and secondly, those related to the learning scenario.

DOTI will be described according to the data type it contains:

- interactions are dyadic; teletandem involves just 2 participants;
- the environment is synchronous (as opposed to non-synchronous such as blogs);
- the time frame is one session (usually from 50 to 60 minutes);
- the communication modality is via VoIP technology;
- communication modes are different such as oral, written via text chat as well as gestures and emoticons.

Specifically, concerning each time frame (i.e. session), the option is given to choose among languages used for communication (e.g. English, Italian) and the number of online sessions (e.g. S1, S2, S3).

Regarding participants, data can be interrogated according to student´s course at the university (e.g. UNESP), gender, and language level (broadly assessed based on performance during teletandem sessions).

In relation to the discourse type, DOTI will be described using free discussion, topic discussion, and task completion (e.g. information/opinion gap).

Metadata for LS are typology of tasks (alternate monolingual interaction or intercomprehension), integrated and non-integrated teletandem modalities (Aranha & Cavalari, 2014), descriptions (aims, materials), teachers' roles, and macrotask and microtask sequences.

DOTI will allow researchers within teletandem contexts to be more coherent in generating, collecting and annotating procedures and thus, will save them time to analyse such multi-faceted, multi-tasking environments more deeply and thoroughly.

Although all the participants have signed consent forms[5] and are enrolled in one of the courses or universities that participate in the Teletandem Network (Leone & Telles, 2016), there are still ethical issues concerned with identification in the future. Hence, we are now considering if the degree of anonymization can be decided on the basis of what participants opt for (i.e. blurring or not their faces).

Besides, a wide range of data is generated every year due to the increasing number of students that participate in the telecollaborative practice. This poses a question of keeping the databank open for including ongoing sessions.

4. Conclusion

For developing criteria of a DOTI, two important concepts have been relevant: interaction space and learning scenario. The former framework places DOTI in a broader field which includes research in corpora compiled by other computer mediated communication such as Facebook or Twitter. Defined metadata will allow us to cross data with other colleagues who are working in the field and there will be guidelines for sharing data collection principles among other colleagues from the teletandem network.

5. The items of the terms vary from institution to institution and an agreement of common ones is still in progress.

DOTI is compiled in an open access corpus perspective. We strongly believe that it will be useful to (applied) linguists, professors, and computer experts who want to develop software based on CMC for language learning.

References

Aranha, S., & Cavalari, S. (2014). A trajetória do projeto Teletandem Brasil: da modalidade institucional não-integrada à institucional integrada. *The ESPecialist, 35*(2), 70-88.

Aranha, S., Luvizari-Murad, L., & Moreno, A. (2015). A criação de um banco de dados para pesquisas sobre aprendizagem via teletandem institucional integrado (TTDII). *(Com) Textos Linguísticos, 9*(12), 274-293.

Brammerts, H. (1996). Tandem language learning via the internet and the International E-Mail tandem network. In D. Little & H. Brammerts (Eds), *A guide to language learning in tandem via the Internet*. Dublin: CLCS.

Chanier, T., Poudat, C., Sagot, B., Antoniadis, G., Wigham, C. R., Hriba, L., Longhi, J., & Seddah, J. (2014). The CoMeRe corpus for French: structuring and annotating heterogeneous CMC genres. *Journal for Language Technology and Computational Linguistics, 2*(29), 1-30.

Foucher, A.-L. (2010). *Didactique des langues-cultures et Tice : scénarios, tâches, interactions*. Université Blaise Pascal - Clermont-Ferrand II.

Leone, P., & Telles, J. (2016). The teletandem network. In T. Lewis & R. O'Dowd (Eds), *Online intercultural exchange: policy, pedagogy, practice* (pp. 243-248). London: Routledge.

Mangenot, F. (2008). La question du scénario de communication dans les interactions pédagogiques en ligne. *Jocair (Journées Communication et Apprentissage Instrumentés en Réseau*, 13-26.

Vassallo, M., &Telles, J. (2006). Foreign language learning in-tandem: teletandem as an alternative proposal in CALLT. *The ESPecialist, 27*(2), 189-212.

Published by Research-publishing.net, not-for-profit association
Dublin, Ireland; Voillans, France, info@research-publishing.net

New directions in telecollaborative research and practice: selected papers from the second conference on telecollaboration in higher education
Edited by Sake Jager, Malgorzata Kurek, and Breffni O'Rourke

Typeset by Research-publishing.net
Cover design and cover photos by © Raphaël Savina (raphael@savina.net)
UNICollab logo – Harriett Cornish, Graphic Designer, KMi, The Open University

ISBN13: 978-1-908416-40-7 (Paperback - Print on demand, black and white)
Print on demand technology is a high-quality, innovative and ecological printing method; with which the book is never 'out of stock' or 'out of print'.

ISBN13: 978-1-908416-41-4 (Ebook, PDF, colour)
ISBN13: 978-1-908416-42-1 (Ebook, EPUB, colour)

Legal deposit, Ireland: The National Library of Ireland, The Library of Trinity College, The Library of the University of Limerick, The Library of Dublin City University, The Library of NUI Cork, The Library of NUI Maynooth, The Library of University College Dublin, The Library of NUI Galway.
Legal deposit, United Kingdom: The British Library.
British Library Cataloguing-in-Publication Data.
A cataloguing record for this book is available from the British Library.
Legal deposit, France: Bibliothèque Nationale de France - Dépôt légal: novembre 2016.