Web-based annotated corpora for language learning: Harnessing their potential
Dr. Catherine Caws
University of Victoria

Date: Oct. 25, 11:30-13:00
Location: RCB 6125

Abstract
Web-based corpora offer many advantages to learners: they show language in use, they
often use authentic material, they can easily be adapted and updated to reflect current
linguistic and cultural norms, hence they provide good evidence of language patterns
(e.g. Bernardini 2004; Braun 2005). Additionally, corpora that are enriched with
annotations, defined by McEnery & Wilson (1997) as “explicit identification of parts of
speech, sentence structure, word meaning, and so on” (10) provide “a great amount of
‘added value’ for the corpus user and consequently extends the potential of corpora in
teaching” (10). Yet, as stated by Braun (2005), corpora are not so often used as main
language resources in language programs; this is often the case because their content has
not been tailored to fit specific learning or instructional goals.

In this paper we will report on the methods used to create a digital library of videos and
annotated transcripts called FrancoToile (http://francotoile.uvic.ca) with the aim of
helping users develop their cultural and linguistic literacies in French. In a view to
enhance learners’ understanding and appreciation of language genres and registers, we
introduced FrancoToile to L2 learners of French who were taking an introduction to
linguistic course. We collected data from users’ interactions, artefacts based on language
tasks, and verbal behaviours. Our data analysis is meant to be further recycled in a re-
design of the systems and of its pedagogical applications.