The use of formulaic sequences by L2 learners has received considerable attention in L2 acquisition research (Dechert, 1983; Hakuta, 1974; Pawley & Syder, 1983; Wray, 2002). Previous L2 studies have mostly focused on two aspects of formulaic language use: comparing L2 speakers’ use of formulaic sequences to L1 norms (De Cock, 2000; Granger, 1998; Warga, 2005) and describing the emergence of formulaic sequences in L2 acquisition (Bardovi-Harlig, 2002; Hakuta, 1974; Myles et al., 1998). Because formulaic sequences are difficult to define and identify, previous research has not justified the choice of target sequences. The identification of formulaic units is often based on native speaker judgments, which adds on more ambiguity when describing their role in L1 and L2 acquisition (Myles et al., 1998). Furthermore, when describing how specific formulaic sequences are acquired, researchers have focused on proficiency level and L1 background as the most important factors that affect the acquisition process. As a result, little attention has been paid to other factors, such as the discourse function of formulaic sequences, which might also affect their acquisition by L2 speakers.

The present study addresses these methodological issues and examines if English speakers’ ability to recall formulaic sequences is affected by their discourse function. This study used corpus-based techniques to identify formulaic sequences that perform different discourse functions in academic discourse (Biber, 2004). Six discourse organizing and six referential formulaic sequences were selected based on their comparative frequencies in the corpus. The target sequences were embedded into an excerpt from an Introduction to Sociology textbook. The text consisted of 13 sections and was recorded by a native speaker of American English. The participants were English native speakers (N = 21) and EFL learners in Russia (N = 43) from two proficiency levels: intermediate and advanced. The participants listened to the recorded text section by section. After they listened to a section twice, they were instructed to do a written reconstruction as close to the original as possible, following the same procedure for every section. The results suggest that there is a main effect for formulaic sequences with higher recall for discourse organizing sequences. The results also indicate that L1 and L2 speakers differ in their recall of formulaic sequences. The implications of this study are discussed in terms of the use of corpus techniques for psycholinguistic research and the role of formulaic sequences in L1 and L2 acquisition research.