Grammaticality and acceptability judgment tasks (GJTs) are found extensively throughout the history of SLA (e.g. Chaudron, 1983 for an early review). These data collection instruments have been used to investigate a range of hypotheses and phenomena, from generativist theories to the effectiveness of instruction. Though popular and convenient, GJTs have led to considerable controversy. There has been concern over their construct validity in efforts to elicit implicit and/or explicit knowledge (Ellis, 2005; Gutierrez, 2013; Ionin & Zysik, 2014). Also contested is the impact of different testing conditions such as untimed vs. timed (and how timing is calculated), mode (oral vs. written), rubrics (with/out correction), target features, and the types of responses required (dichotomous vs. scaled; with/out an indication of the basis and/or confidence of the judgment).

This paper presents a synthesis of research involving judgment data. Based on a comprehensive search, including a corpus of approximately 8000 journal articles published 1980-2015, we code the use of GJTs for their different methodological and psychometric properties, as well as the nature of the theories, constructs, and questions under investigation. We discuss trends of usage over time, such as combinations with other techniques and data analyses employed. We also consider the potential role of methodological transparency by examining a subset of approximately 60 GJTs held on IRIS (www.iris-database.org).

This paper builds on a small but growing body of methodological reviews in second language research that focus on data collection procedures, such as Derrick’s (2016) and Plonsky & Derrick’s (in press) reviews of instrument reporting practices and reliability coefficients, respectively. Though our focus is on research previously conducted, we use our results to look ahead and provide the field with empirical evidence for understanding and improving research involving GJTs and similar techniques.

Keywords: GJT, research synthesis, research methods, data collection.