Problems of a real-time sociolinguistic project: Lessons from the Okazaki Survey on Honorifics

KENJIRO MATSUDA
Kobe Shoin Women's University

With the rise in real-time data, the past 20 years has seen the emergence and flourishing of sociolinguistic panel data analysis. Access to the kind of data that sociolinguists had long dreamt of did not betray their expectations, and led to a number of new findings, including the late *adoption phenomenon* in Montreal English (Boberg 2004) and possible sex differentiation in that phenomenon (Sankoff and Blondeau 2007), among others. In exchange for the new data, however, sociolinguists now must face a new series of problems specific to panel data. Drawing on the experience of the Okazaki Survey on Honorifics (OSH), a real-time questionnaire-based sociolinguistic project studying changes in the use of and consciousness surrounding honorifics in the city of Okazaki, Japan (Matsuda 2012), I pick up some such issues and make some proposals regarding them.

Many of the problems stem from the fact that real-time data involves decades of time lapse between two data-collection points. In the case of OSH, the interval was 55 years, which brought about three major inconveniences. 1. Massive socio-cultural changes made the elicitation pictures virtually unusable 2. Later developments in linguistics (especially in discourse-related fields) led to the identification of several inadequacies in the questionnaire. 3. The decreasing sample size has made it impossible to analyze the responses in a statistically reliable way.

No less serious than these problems is the issue of a proper statistical model with which to analyze the panel data. Since panel data involves repeated observations, the familiar generalized mixed-effects model cannot be used. In light of the fact that one needs to separate the effects of age, period, and cohort (Glenn 2005), the Nakamura Bayesian Cohort Model (Nakamura 1986) emerges as the most proper statistical model.