Worldlikeness: A Web-based tool for typological psycholinguistic research

Tsung-Ying Chen and James Myers
Graduate Institute of Linguistics, National Chung Cheng University, Taiwan

Web-based tools for running psychological and psycholinguistic experiments have become more prevalent over the last decade, since they allow researchers to crowdsource data from a large number of speakers (e.g. *tatool*: von Bastian et al. 2013; *turktools*: Erlewin & Kotek, in press). Typological research, however, must not only test many speakers, but many languages, while maintaining methodological consistency. In this talk, we introduce WORLDLIKENESS, a Web-based tool (currently hosted on an Amazon cloud server at http://lngproc-4083.nitrouspro.com:3000) for collecting and sharing cross-linguistic wordlikeness judgments, and demonstrate its potential with small-scale studies on Mandarin, Taiwan Southern Min, and English.

Web-based experiments have played an important role in the growing use of megastudies in psycholinguistics (Keuleers & Balota 2015). Megastudies use very large samples so that regression-based designs can tease apart lexical variables, which tend to be both gradient and partially confounded. For example, wordlikeness judgments are affected both by phonotactic probability and neighborhood density; even though these variables are conceptually quite distinct, they are highly correlated, leading Bailey & Hahn (2001) to distinguish them through regression analyses of judgments for a randomly generated set of nonwords. Linguistic variables are partially confounded across languages as well (e.g. languages with smaller syllable inventories, like Mandarin, tend to have higher neighborhood densities), but extending the megastudy approach across languages faces logistical challenges: Typological psycholinguists need to test many speakers from each language, using materials designed with the help of expert native speakers (e.g. Bates et al. 2003 required 22 coauthors just to test seven languages). Worldlikeness aims to make typological psycholinguistics feasible, by providing tools for separate groups of experimenters to design materials for individual languages, collect judgments online, and crucially, share their data with each other for typological analysis.

To the best of our knowledge, Worldlikeness is the first Web-based experiment tool that integrates three subsystems: (i) an experimenter subsystem, with a user-friendly interface for uploading old results to share and for designing new experiments with a variety of response scales and trial structures, and stimuli (written, auditory, or video) that can be uploaded or (for written stimuli) automatically generated (along with their phonotactic probabilities and neighborhood densities) to maintain cross-study consistency, (ii) a participant subsystem, which allows for a customizable native-speaker test to help screen participants, runs experiments on participants' computers or mobile devices, collects key presses (or clicks via mouse or touchscreen) and response times, passes them to the experimenter with participant-, item-, and trial-level information, and rewards participants with a colorful and concise report comparing personal statistics with

Figure 1. A result report seen by the participant after completing an experimental session

group results (Figure 1), (iii) a researcher subsystem, which allows registered experimenters or the general public to search for experiments and results in the online archives, for downloading and typological analysis. The language of the Worldlikeness interface can also be changed to match that of the experimenters or participants. Worldlikeness is designed to follow strict research ethics

guidelines. A password-protected account system helps experimenters manage (add, publish, privatize, and remove) their experiments and results, with private results secured on the cloud server. Participants maintain their anonymity throughout (attempts to retake an experiment are blocked by checking their publicly available IP address). As part of the customizable online consent form, participants set the access level for their own data (restricted to the experimenter only, available only to registered experimenters, or freely available to the general public). Results that experimenters have shared are automatically filtered in accordance with participants' privacy preferences. Moreover, since Worldlikeness is built as an open-source Web application in Meteor (www.meteor.com), it can run in any modern browser without any additional software. It is not only free of charge, but open-source, so Web developers are welcome to extend our work.

Although Worldlikeness is less than a year old, we have already used it to run small-scale replications of the Mandarin wordlikeness megastudy reported in Myers (2015). In the original study, over 3,000 randomly generated nonlexical monosyllables, divided into two random sets, were presented in Zhuyin Fuhao (the phonetic orthography used in Taiwan) to over 100 Mandarin speakers, with good/bad judgments and reaction times collected using E-Prime (Schneider et al. 2002). In the Worldlikeness replications, novice experimenters randomly selected a subset of 100 items from each of the two sets, and asked two groups of 10 speakers to judge each subset. Both replications showed significant correlations with the original judgments in both by-item mean acceptability and response time (Figure 2).
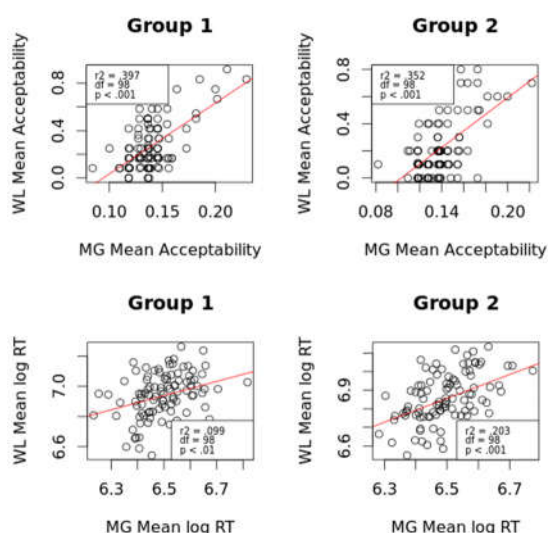

Figure 2. Correlation between mean acceptability and log-transformed response time in the Myers (2015) megastudy (MG) and the Wordlikeness (WL) replications

We are currently using Worldlikeness to compare these Mandarin results with those from Taiwan Southern Min (a related language with a somewhat larger syllable inventory) and English (with a much larger syllable inventory). Data come both from old data uploaded to Worldlikeness (Myers & Tsay 2012 for Southern Min, Albright & Hayes 2003 for English) and from a new Worldlikeness experiment testing Mandarin and Southern Min bilinguals on randomly generated syllables that are nonlexical in both languages. We are eager to invite more linguists to join us in Worldlikeness, to extend the database to a wider variety of languages and help make truly typological psycholinguistics a reality.

References **Albright & Hayes** (2003). Rules vs. analogy in English past tenses. *Cognition, 90.* **Bailey & Hahn** (2001). Determinants of wordlikeness: phonotactics or lexical neighborhoods? *JML, 44.* **Bates et al.** (2003). Timed picture naming in seven languages. *Psychonomic Bulletin & Review, 10.* **Erlewin & Kotek** (in press). A streamlined approach to online linguistic servers. *NLLT.* **Keuleers & Balota** (2015). Megastudies, crowdsourcing, and large datasets in psycholinguistics. *QJEP, 68.* **Myers** (2015). Markedness and lexical typicality in Mandarin acceptability judgments. *Language & Linguistics, 16.* **Myers & Tsay** (2012). The interaction of markedness and experience in phonotactics judgments. *SLE, 45.* **Schneider et al.** (2002). *E-Prime: User's Guide.* PSI. **von Bastian et al.** (2013). Tatool: A Java-based open-source programming framework for psychological studies. *Behavior Research Methods, 45.*