

# Crowdsourcing a valid option for gathering speech ratings

February 18 2015

---



Credit: AI-generated image ([disclaimer](#))

Crowdsourcing – where responses to a task are aggregated across a large number of individuals recruited online – can be an effective tool for rating sounds in speech disorders research, according to a study by NYU's Steinhardt School of Culture, Education, and Human Development.

"Because large crowdsourced samples can be obtained quickly, easily, and inexpensively, [speech](#) researchers could find it beneficial to use crowdsourcing technology in place of traditional methods of collecting speech ratings," said Tara McAllister Byun, an assistant professor in NYU Steinhardt's Department of Communicative Sciences and Disorders and the study's lead author.

Research in linguistics and psychology has reported that using crowdsourcing not only saves time and money, but can actually enhance scientific rigor. The NYU study, published in the *Journal of Communication Disorders*, suggests that these benefits can also be extended to studies of the nature and treatment of [speech disorders](#).

In speech disorders research, unbiased [listeners](#) are needed to evaluate patients' progress over the course of treatment by listening to [speech sounds](#) and rating or coding them. Because [speech language pathologists](#) and other trained professionals are often used as raters, collecting the ratings can be costly. It can also be a challenge to find raters who are not part of the research and are therefore unbiased.

Amazon Mechanical Turk (AMT) is an online crowdsourcing platform developed by Amazon as a tool for completing routine tasks better performed by humans than computers. Now with hundreds of thousands of workers, and roughly 10,000 requestors or employers, anyone can use AMT's standardized interface to post or complete electronic tasks. While not originally designed for conducting behavioral research, AMT has been successfully used in linguistics and psychology research.

Modeling studies have shown that even when individual responses to a task are not highly accurate, aggregated or crowdsourced responses from a large number of people generally converge with those of experts. In this study, the researchers tested the validity of having AMT users rate speech sounds, compared with ratings collected from experienced

listeners.

Listeners were asked to rate recordings of 100 words containing the "r" sound, collected from children with trouble pronouncing the sound and working to correct it in speech therapy. Twenty-five experienced listeners and 153 AMT listeners scored the "r" sounds as correct or incorrect. Data from experienced listeners were collected over a period of three months, while data gathering using AMT took a mere 23 hours.

The researchers found that when responses were aggregated, there was a very high level of overall agreement. When items were classified as correct or incorrect based on the majority vote across all listeners in a group, the AMT group and the experienced listener group were in agreement on all but seven of 100 items.

In a further analysis, the researchers sought to understand how many AMT listeners were needed to still get valid responses that converged with those of experienced listeners. They found that samples of nine or more AMT listeners demonstrate a level of performance consistent with typical expectations for experienced listeners.

While using AMT for speech ratings poses some limitations, including a lack of control over sound quality and inattentive or uncooperative raters, the researchers concluded that using AMT for speech language pathology research could have a substantial impact on the process of gathering speech ratings.

"A key advantage of using crowdsourcing to recruit listeners for speech rating tasks is the speed and ease with which ratings can be obtained," said McAllister Byun. "However, using [crowdsourcing](#) for speech data rating is not merely a question of convenience; it also has the potential to improve speech research by expanding access to independent listeners, thereby reducing bias."

Provided by New York University

Citation: Crowdsourcing a valid option for gathering speech ratings (2015, February 18)  
retrieved 25 April 2024 from

<https://medicalxpress.com/news/2015-02-crowdsourcing-valid-option-speech.html>

This document is subject to copyright. Apart from any fair dealing for the purpose of private study or research, no part may be reproduced without the written permission. The content is provided for information purposes only.