Can we use the internet to study speech production? Yes we can! Evidence contrasting online versus laboratory naming latencies and errors

Amie Fairs (Aix-Marseille Université) & Kristof Strijkers (Aix-Marseille Université)
Amie.fairs@univ-amu.fr

The closure of cognitive psychology labs around the world due to the COVID-19 pandemic has prevented in-person testing. This has caused a particular challenge for research involving speech production, as before the pandemic there were no studies demonstrating that reliable overt speech production naming latencies could be collected via the internet. In fact, concerns over the reliability of timing between picture presentation and the onset of audio recording of participants responses led many researchers to at least informally have a negative view of online testing. Here, we present evidence that both accurate and reliable timings of overt articulation data can be collected from internet-based speech production experiments.

We tested 100 participants in an internet-based picture naming paradigm, where we orthogonally manipulated the word and phonotactic frequency of the picture names, with 55 target items per cell. We compared our results to a lab-based study [1] which used the same materials and design, giving us a benchmark to test our results against. In the lab-based study, we found a significant effect of word frequency but no effect of phonotactic frequency, and our main question was to determine if the same pattern of results would be found when running this study via the internet. Aside from this main question, we additionally investigated 1) whether participants' internet upload or download speed would have an effect on their production latencies or errors, and 2) planned a permutation-style analysis on our data to determine how many participants were necessary to detect a significant effect. Data were analysed using (generalized) linear mixed effects models.

Regarding our first question, we found a significant word frequency effect but no phonotactic frequency effect, fully replicating the lab-based results. The raw sizes of the word frequency (51ms vs 58ms) and phonotactic frequency (9ms vs 1ms) effects were similar between the internet-based and lab-based studies. However, the production latency distribution was different, with significantly longer latencies by 100ms in the internet-collected data compared to the lab data (see Figure 1 for latency distributions of the internet-based and lab-based studies). Regarding our second question, we found no substantial evidence that internet upload or download speed affected either naming latencies or errors. Regarding our third question, we randomly selected different sample sizes of participants and tested these smaller sets for significant word and phonotactic frequency effects, repeating this process 1000 times per sample size. The results suggest, using a conservative threshold, a minimum sample size of 40 participants for online production paradigms.

In sum, we demonstrate that internet-based testing of speech production is a feasible and promising endeavour, where robust lab effects, such as the word frequency effect, are also detected with a similar effect size when tested via the internet. Participants' internet speed did not appear to affect either production latencies nor error proportions, which was informally thought to be a big barrier to accurate implementation of online studies. However, we advise future production researchers to collect internet speed information from participants. Overall, collecting overt speech production data via the internet has less challenges than many researchers (anecdotally) assumed.

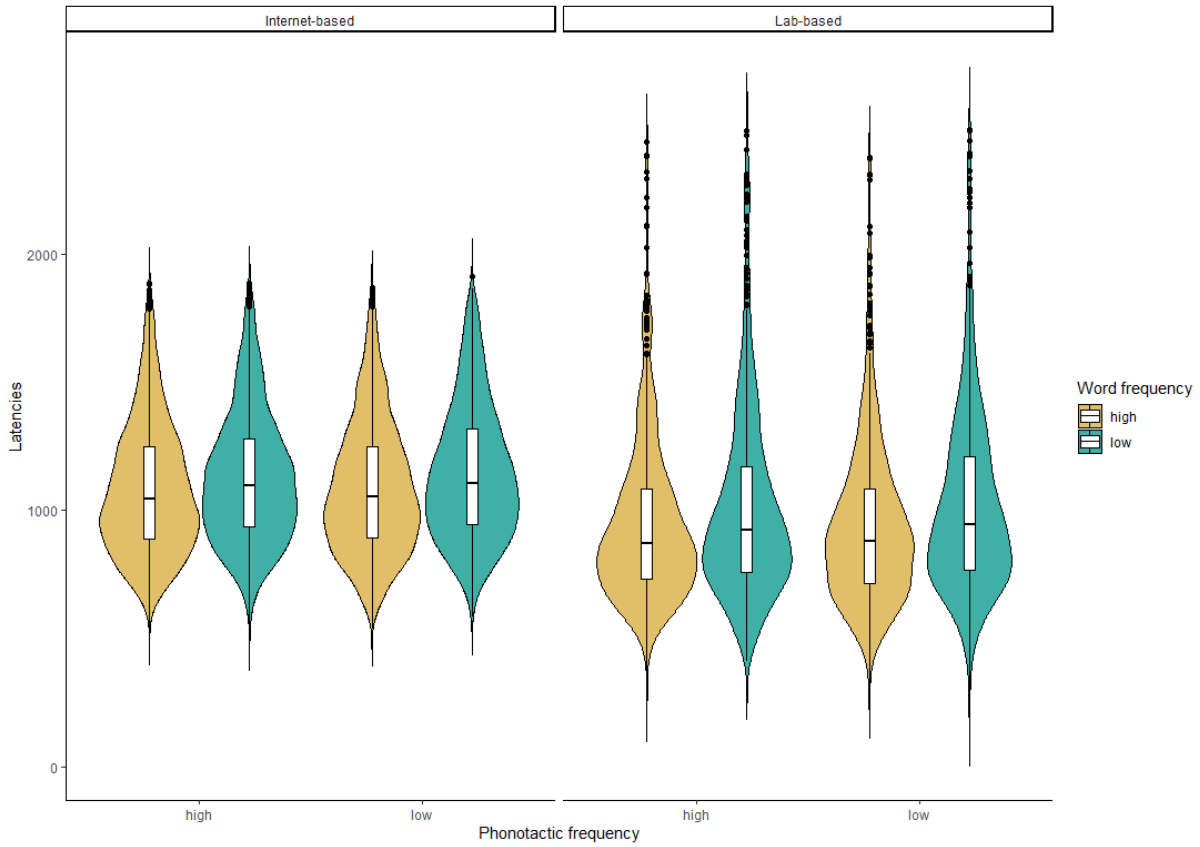

Figure 1. Latency distributions of the internet-based experiment data (left panel) and lab-based experiment data (right panel). The x-axis plots the phonotactic frequency manipulation, with the left two beans in each panel representing high phonotactic frequency items and the right two beans representing low phonotactic frequency items. Yellow beans represent high word frequency items and green beans represent low word frequency items.

References

[1] Fairs, A., Michelas, A., Dufour, S. & Strijkers, K. (2021). The same ultra-rapid parallel brain dynamics underpin the production and perception of speech. *Cerebral Cortex Communications,* 2:3, tgab040, https://doi.org/10.1093/texcom/tgab040.