DE GRUYTER MOUTON Linguistics Vanguard 2024; aop

Claire Nance*, Maya Dewhurst, Lois Fairclough, Pamela Forster, Sam Kirkham, Justin J. H. Lo, Jessica McMonagle, Takayuki Nagamine, Seren Parkman, Haleema Rabani, Andrea Siem, Danielle Turton and Di Wang

Articulatory phonetics in the market: combining public engagement with ultrasound data collection

https://doi.org/10.1515/lingvan-2024-0020
Received January 29, 2024; accepted September 30, 2024; published online December 19, 2024

Abstract: This paper is a methodological contribution outlining an approach we have developed to recording ultrasound tongue imaging and audio research data as part of public engagement work. The paper is based on our experience of recording in East Lancashire, UK, at two such events as well as building on the work of other colleagues who have conducted similar projects. We have successfully managed to collect relevant articulatory research data while providing an interesting and enjoyable event for the public. In a context of stretched research budgets and researcher time constraints, this combination allows a time-effective combination of tasks, whilst democratizing academic work and engaging local communities. Our paper has two aims: (i) to describe the logistical and ethical considerations for organizing an event combining articulatory research and public engagement, and (ii) to provide methodological reflection on data collection and eventual data quality obtained including assessment of background noise. We hope to provide inspiration and suggestions for colleagues wishing to pursue research in this area and also acknowledge where some of our methods would benefit from more effective solutions. Sample documents for ethics, publicity, risk assessments, staff planning, timelines, and budgeting are included in our supplementary material toolkit available at https://osf.io/ky3cz/.

Keywords: public engagement; ultrasound tongue imaging; phonetics; sociolinguistics; data quality

1 Introduction

Recent years have seen the expansion of data collection methods in speech production and sociophonetic research beyond traditional techniques. Some of these methods, such as remote recording, were necessitated by COVID-19 restrictions (Broś 2024; Freeman and De Decker 2021; Kostadinova and Gardner 2024; Leemann et al. 2020; Sekerina et al. 2024; Sevilla 2024; Zhang et al. 2021). Remote crowdsourcing of speech data has also been employed in order to leverage larger sample sizes than would be possible in person (Kim et al. 2019; Love et al. 2022) or to adopt a citizen science approach to obtaining audio data (Koreinik et al. 2024). Similarly, researchers have exploited the near-ubiquity of smartphones in many societies for accessing large sample sizes (Kirkham et al. 2020; Leemann et al. 2018; Strycharczuk et al. 2020), and engaging minority language users in revitalization programmes (Hilton 2021).

In this paper we present our approach to a specific kind of data collection: combining articulatory phonetics and public engagement work outside the lab. While it has long been common practice to collect articulatory phonetic data outside universities in order to access populations who are geographically distant from the laboratory (for example, Ladefoged 1968, 2003; Whalen and McDonough 2015), we focus particularly on how ultrasound tongue imaging data can be combined with a programme of public engagement (Heyne et al. 2020; Smith et al. 2023; Strycharczuk et al. 2023).

***Corresponding author: Claire Nance**, Lancaster University, Lancaster, UK, E-mail: c.nance@lancaster.ac.uk. https://orcid.org/0000-0001-5953-155X
Maya Dewhurst, Lois Fairclough, Pamela Forster, Sam Kirkham, Justin J. H. Lo, Jessica McMonagle, Takayuki Nagamine, Seren Parkman, Haleema Rabani, Andrea Siem, Danielle Turton and Di Wang, Lancaster University, Lancaster, UK. https://orcid.org/0000-0003-0115-6982 (J.J.H. Lo)

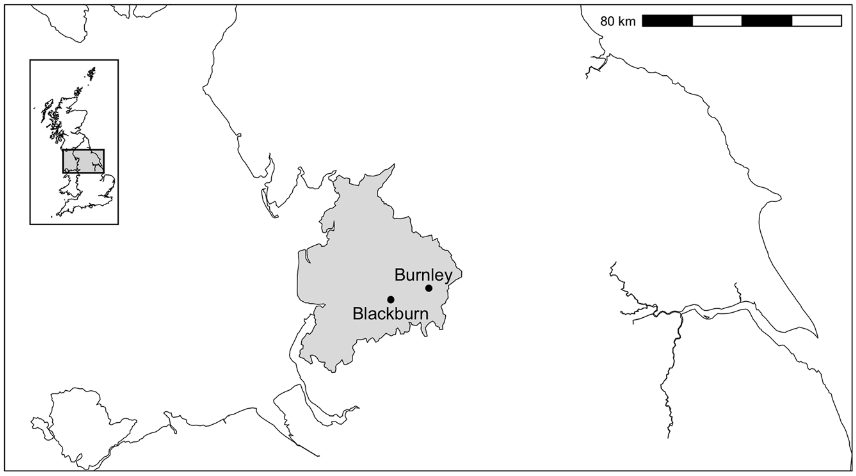

Figure 1: A map showing the locations of Blackburn and Burnley within the north-west of England. Grey shaded area shows the ceremonial county of Lancashire.

We recorded ultrasound data in covered market halls from 32 participants in Blackburn in 2022 (initial findings reported in Nance et al. 2023), and from 32 participants in Burnley in 2023. Both locations are large towns in East Lancashire, in north-west England (see Figure 1). Our approach involves setting up research equipment outwith a lab setting. Another approach which could be extremely fruitful (if budgets allow) is building a mobile lab and taking this directly to the public (Tiede et al. 2024; Wieling et al. 2023).

In a context where research funding as well as researcher time are in short supply, we outline how engagement, research, and postgraduate student development can be combined in an effective manner. At the same time, these methods contribute to a mission of sharing and democratizing access to academic research, as well as providing the opportunity for the academic community to learn from members of the public (Price and McIntyre 2023; Rymes and Leone 2014). When designing the two events we have run so far, we have primarily aimed to provide an engaging and interesting event for the public. At the same time, we also aimed to collect some research data and provide a structured framework for postgraduate students to conduct public engagement work. We believe it necessary to order the aims in this way to make public engagement as effective as possible.

This paper has two aims:

(1) We first describe the logistical and organizational considerations involved in organizing and running this kind of event (Section 2).
(2) We then discuss equipment and methodological considerations for research designs which work well in this setting, and the quality of the eventual data obtained including assessment of background noise (Section 3).

Our paper is supported by a resource toolkit containing research materials, sample ethics documents, risk assessments, publicity materials, step-by-step planning timeline, equipment lists, and sample budget (available at https://osf.io/ky3cz/). These materials and methods were designed for the UK context, but we hope that they could provide a framework and ideas for use in other contexts.

2 Logistics and organization

In this section we address logistical considerations including budget and timeline, publicity, ethics, safety, and staffing, and organization before and after the event.

2.1 Funding and timescale

We have previously obtained funding from two different public engagement festivals in the UK: the Festival of Social Sciences (https://festivalofsocialscience.com/) in 2022 and Being Human Festival (https://www.

beinghumanfestival.org) in 2023. It would be possible to do something like this independently with other funding sources, but we found it useful to have the support and training of a wider national festival. Our events worked to bring people into the market for something different from their usual weekend activities, which also benefited the market due to increased visitor numbers. Hosting events in market halls has the advantage of existing footfall, so potential participants do not necessarily need to know about the event in advance. Markets also typically offer existing infrastructure such as good transport links, toilets, food, and an accessible building.

The budget needed for this kind of work is quite low. We have previously worked with £1,200–£2,000. From this, the greatest expense is properly paying postgraduate students to contribute to the work. Another large budget component is publicity materials (see Section 2.2). In both cases we ran our events as one-off pop-up stalls, and we were offered the market stall and power supply for free. Depending on the location, it might be necessary to hire a stall or similar space for the day. Superficially, it could seem very attractive to hold a one-day event and record research data from 30+ participants. However, in each case, preparation for these events has taken up to 9 months. We have included a step-by-step timeline in the supplementary materials to indicate all the tasks involved in organization of these events.

2.2 Publicity

To recruit as many participants as possible, we conducted extensive publicity ahead of our events. We set out to develop a publicity and marketing strategy aiming to recruit our target audience (local people who might not necessarily otherwise engage with university research) and engage them once they arrived at our market stall. We carefully considered our target audience and the kinds of media they were likely to engage with. We then worked with our university press office to send a press release to local and regional newspapers, websites, radio, and TV. The press release was especially taken up when a local interest “angle” was included: we focused on a sporting rivalry, football, between Blackburn and Burnley to generate interest and then explained that we were analysing accent differences between the locations. We also ran geographically targeted advertisements on social media in the week ahead of our event, posted in local groups, and posted on a regional events website. Finally, we asked staff and students with connections in the area to advertise among friends and family.

2.3 Preparing for the event

We obtained ethical approval from our university for recording our data. Adult participants read an information sheet, then signed a consent form. For child participants, we asked them to sign an assent form and asked their guardian to sign the consent form. After the recording, we collected a small amount of biographical information from each participant via a paper questionnaire. Insurance was covered by our university’s public liability insurance including third party liability, and departmental insurance for our lab equipment. Risk assessments were carried out by the first author and shared with our venue. While working in a public space we carefully considered the safety of both our team and the participants. We also discussed informed consent, and working with people who might not be able to give informed consent. We cleaned all ultrasound equipment between participants. From the point of view of keeping our team safe, we stored the phone number for the market managers and were in contact with the market security teams. On the morning of each event, we received a briefing for fire and safety procedures.

We recruited seven students to work as part of our event in Burnley. At the market stall, we made up a team of four staff members and six paid students. The additional student worked on publicity ahead of the event and data post-processing. Each student was paid for 10 h, made up of 2 h of training and 8 h of working on the market stall, travel, set-up, and take-down. We would recommend training students in organization, safety, accessibility, and taking consent from participants, as well as in using the ultrasound equipment. We selected a team who had the communication skills to work well with the public, and who represented a range of experience and backgrounds. We took quite a large team with the view that some spare capacity is needed. This allowed us all to take adequate breaks. Our budget allowed all team members some money for food and drinks.

2.4 During the event

We designed the set-up of our stall to include a clear participant workflow including:
(1) Welcome and demonstration
(2) Ethics
(3) Recording
(4) Background questionnaire, postcards (see Section 3.2), participant gifts
(5) Evaluation

Each participant was allocated a number when they arrived at the ethics station. This was written on their documents, and their data stored under this number. We discussed inclusive practice and adaptations at the planning stage of our project. We successfully adapted the activity for visually impaired participants and had a plan in place if deaf participants wished to take part. For visually impaired participants we read out the information sheet and consent form while taking consent. We then described the target stimuli and spelled them out before participants produced the word. For deaf and hard of hearing participants, we planned to allow them time to read our information sheets and publicity materials, and then demonstrate the equipment and recording before taking consent. Ideally, we would have included a colleague with British Sign Language skills as part of our team. We also worked with participants with cerebral palsy, learning disabilities, and mobility disabilities.

We displayed a notice that photography would be taking place and asked any adults whose face appeared prominently in photos to sign a photo marketing permissions form from our university. We asked guardians of any children appearing in photos to sign the permissions form, but in general avoided taking pictures of children. We also hosted a film crew from local BBC TV news during our event in Burnley. The programme was broadcast on the evening of our event so could not contribute to recruitment on the day, but we believe it raised the profile of our work and will help with planning for future events. Also, the programme provided some positive publicity for our hosts at Burnley Market. Where possible, we asked participants to scan a QR code with a link to the funder's evaluation survey as they were leaving the stall. Ideally, we would recommend trying to accurately record the exact footfall (as well as research participants recorded).

2.5 After the event

On an opt-in basis, we asked participants if they would like to receive a sample video of their tongue after the event. Most of the participants opted in and we asked them to leave an email address. We then exported a full-speed and slow-motion video of one word from their data and emailed these out. At the same time, we offered participants the chance to stay in touch and be informed about project results. We emailed our participants videos to provide a parallel experience to ultrasound pictures of babies from pregnancy scans. Another option would be to give participants a print-out image of their tongue (Wieling et al. 2023). Working with the public in this way creates a kind of social contract (Golumbic et al. 2017; Svendsen 2018). It is therefore important for us to continue the reciprocal nature of our relationship and offer accessible reports of our results. For this reason, we have offered to give talks in public venues where we have recorded and will send lay-person summaries of the findings, as well as research articles, to interested participants.

3 Methods and data

In this section we discuss our research design considerations and recording equipment, and we assess the quality of the data collected.

3.1 Research design

Our research questions concern (i) the extent of coda rhoticity across different locations and sectors of the population across East Lancashire, and (ii) how coda rhoticity is realized in articulatory terms. Coda rhoticity (pronouncing rhotics in words like *far* and *farm*) is variable across East Lancashire and declining in apparent time (Barras 2010; Dann et al. 2022; Ryan et al. 2022; Turton and Lennon 2023). We chose to record data through public engagement in Blackburn and Burnley as we thought we would be able to get a relatively geographically homogenous sample of participants necessary for these research questions. Other studies taking our methodological approach have instead opted to address questions which can be answered with participants from broader dialect areas such as the south of England (Strycharczuk et al. 2023), New Zealand (Heyne et al. 2020), or “English speakers” (Smith et al. 2023; Tiede et al. 2024). These larger dialect specifications allow greater scope for using data from all interested participants.

When designing our study, we thought carefully about the kind of ultrasound data that can be imaged in midsagittal view and would likely capture a reasonable range of variation between the speakers we could record. It is perhaps no coincidence that the majority of studies conducted so far using ultrasound recordings and public engagement have all considered aspects of liquid consonants (Heyne et al. 2020; Nance et al. 2023; Strycharczuk et al. 2023; Tiede et al. 2024). All these studies have selected liquids as easily accessible with ultrasound tongue imaging, and dialectally variable in English. Smith et al. (2023) instead investigate vowel production consistency.

3.2 Data recording and equipment

Our data recording set-up included two ultrasound machines and laptops so we could record two people simultaneously and independently (see Figure 2). This set-up has the additional benefit of providing some redundancy in case of technical issues. We recorded simultaneous audio and ultrasound data in Articulate Assistant Advanced (AAA; Articulate Instruments 2022). The ultrasound data were recorded using a Telemed MicrUs ultrasound machine, with a 64-element 20 mm radius probe. The frame rate was approximately 80 fps, probe frequency 2 MHz, depth 80 mm, and field of view 90–101 %. Participants wore an UltraFit headset for probe stabilization under the chin (Spreafico et al. 2018). The audio data were recorded with a Beyerdynamic Opus 55 condenser microphone attached to the ultrasound headset at 22,050 Hz sampling rate directly into AAA via a Sound Devices USB Pre-2 audio interface. The audio and ultrasound data were synchronized by the AAA software with an Articulate Instruments Pulse Stretch Unit.

Before reading the word list, participants were first recorded pressing their tongue against a plastic bite plate. Ultrasound images are rotated to this image of the occlusal plane to allow better cross-speaker comparison

Figure 2: Our recording set-up with two ultrasound stations in Burnley Market. Justin J. H. Lo and Takayuki Nagamine are recording participants.

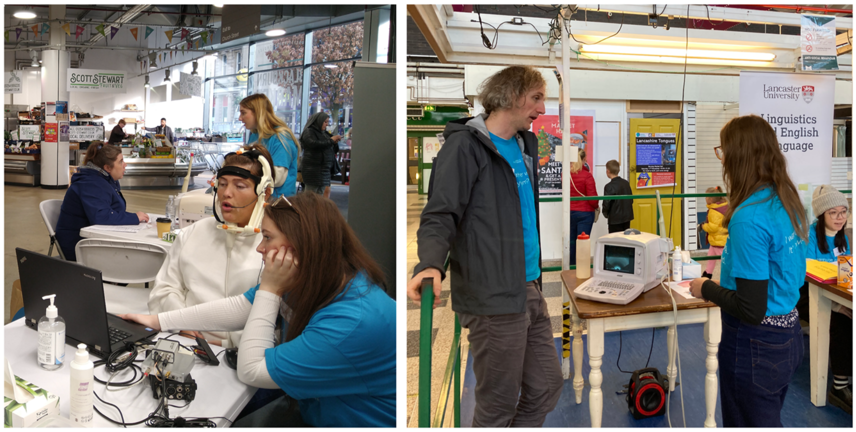

Figure 3: *Left*, Maya Dewhurst is recording data from a participant in Blackburn; behind them, Lois Fairclough is using the Mindray DP-2200 machine to demonstrate ultrasound to a participant. *Right*, Sam Kirkham and Seren Parkman are preparing the welcome station in Burnley.

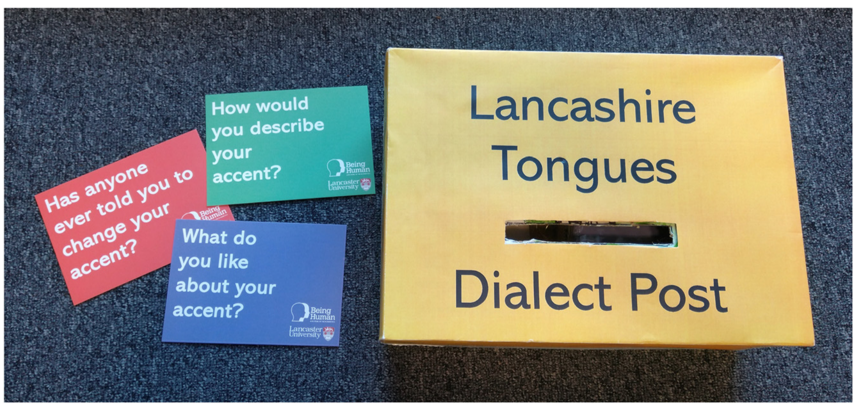

Figure 4: Accent attitude postcards and our "postbox".

(Scobbie et al. 2011). We then made a recording of each participant swallowing some water. Filling the oral cavity with water usually provides a good image of the hard palate, which can then be used for future analysis or plotting. For further guidance on best practices for recording and analysing ultrasound data, see Lawson and Dokovova (2023) and Balch-Tomes and Wrench (2024).

We recorded data from anyone who was interested. We then analysed the data relevant to our research questions. For an evaluation of the data we obtained, see Section 3.3. To keep the event fun and within a reasonable time constraint, we opted to use a short word list of 13 words in isolation with two repetitions. The words were presented with pictures as well as orthography to make them as accessible as possible. We included extra components to the event to try and make it interesting for participants in addition to the ultrasound recording, for example the dialect postcards described below. In the area where people were likely to arrive at the stall, we had a welcome station with a third ultrasound machine (Mindray DP-2200). This machine was used to demonstrate the technology to potentially interested people and show them what we were doing (Figure 3).

After recording, we invited participants to share a small amount of biographical information via a questionnaire. We also had postcards they could optionally fill in, each one containing a simple question about their opinions on accents. Participants could write on the back of these anonymously and "post" the card to us in a box for this purpose (see Figure 4). We did not intend for this activity, adapted from the Dialect and Heritage Project (https://dialectandheritage.org.uk/), to be used as research data but it could be used as such in the future.

3.3 Assessment of the data collected

In this section we consider the quality of data collected. We aimed for participants to complete the whole process of giving consent, recording, and background questionnaire within 15 min, though some chose to stay longer. This data collection method allowed us to have a relatively large participant sample for an articulatory phonetic study, though with a relatively small token count per participant. These factors must be balanced according to the research aims of a particular project and the statistical modelling strategy.

The quality of ultrasound data obtained can vary depending on the skill of the researcher in fitting the headset (Pucher et al. 2020) and individual participant anatomical factors such as chin size, facial hair, or a dry mouth (Scobbie 2013). As such, it is not meaningful to compare ultrasound datasets across different data collection sites as it would not be possible to obtain fully comparable data, and site is not the greatest factor affecting quality. We excluded three speakers from our analysis whose ultrasound images were not of sufficient quality to allow accurate spline fitting to take place (out of 44 Blackburn and Burnley participants eligible for research analysis). This is comparable to previous ultrasound datasets we have collected. It is possible that participants felt more self-conscious about being recorded in a public setting rather than in a lab, but it is also possible that they felt more comfortable being recorded in a setting they knew. There are no specific limitations on the kinds of analysis which could be conducted on the data. For example, Strycharczuk et al. (2023) have done detailed analysis on the distance between the short tendon and tongue tip, and Smith et al. (2023) use their data as a control sample for comparison with a clinical population.

In the following subsections, we assess the acoustic data in more detail, especially focusing on background noise and potential effects on acoustic measurement accuracy (Section 3.3.1). We then discuss the characteristics of the participants we recorded (Section 3.3.2).

3.3.1 Evaluating acoustic data background noise

Preliminary data processing has allowed us to quantitatively assess the audio obtained from recording in market settings. Here, we focus on the extent of background noise as a potential source of inaccurate acoustic measures in data analysis. As data collection methods outwith lab settings have proliferated in phonetics, so too have analyses focused on evaluating the accuracy of acoustic measures derived (Conklin 2023; Freeman and De Decker 2021; Kostadinova and Gardner 2024; Sanker et al. 2021; Sekerina et al. 2024; Sevilla 2024). However, as noted by Conklin (2023: 2), much of this research concentrates on comparing software and hardware, rather than background noise evaluation. Typically, background noise in remotely collected data is evaluated qualitatively, and potentially problematic recordings manually excluded (Kim et al. 2019: 161; Kirkham et al. 2020: EL73; Strycharczuk et al. 2020: 5).

Here, we assess signal to noise ratio (SNR) as a measure for quantifying the background noise compared to the speech signal recorded (De Decker 2016; Maryn et al. 2017; Sanker et al. 2021). SNR (in decibels) is calculated as:

$$20\, log_{10}\ \frac{signal}{noise}$$

where “signal” is the maximum amplitude (in pascals) of the speech part of a recording compared to the maximum amplitude of a section of recording where the participant is silent (“noise”).[1] Higher numbers indicate less noisy recordings. For high-quality recordings suitable for gold-standard use in clinical diagnosis, Deliyski et al. (2005: 27) recommend a SNR of at least 30 dB for accurate measurements. Other estimates are less conservative: for example, recordings with SNR greater than 10 dB are recommended for accurate jitter and shimmer measures (Ingrisano et al. 1998: 94), and greater than 12 dB is recommended for clinical recordings in Titze (1995: 29). Mid-spectrum frequencies (1,000–2,700 Hz) are most likely to be adversely affected by background noise and produce inaccurate acoustic analysis (Parikh and Loizou 2005: 3882). Similarly, De Decker (2016: 17) demonstrates that acoustic measures of high vowels, F2, and from recordings of male speakers are most likely to be inaccurate with background noise. De Decker (2016: 17) suggests that recordings with SNR less than 10 dB should be considered “noisy” and likely to produce inaccurate acoustic measures.

1 Equation recommended by the US National Institute of Standards and Technology at https://www.nist.gov/itl/iad/mig/nist-speech-signal-noise-ratio-measurements (accessed 21 November 2024).

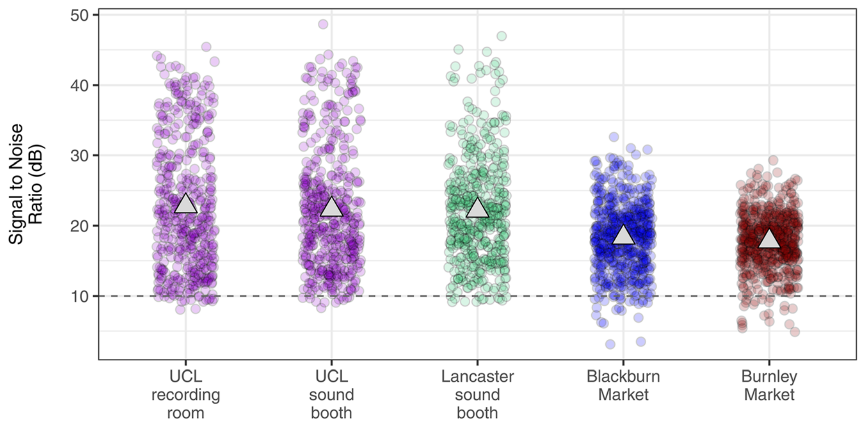

Figure 5: Signal to noise ratios for recordings conducted at UCL (recording room and sound booth), Lancaster, Blackburn, and Burnley. Grey triangles show the mean value for each context. The dashed line at 10 dB shows the threshold for "noisy" data defined in De Decker (2016).

In our marketplace recordings, the data from eligible participants comprise 571 tokens from Blackburn, and 601 from Burnley. We calculated SNR using a Praat script with the default sinc70 interpolation method (Boersma and Weenik 2023). For the "noise" comparison to speech, we measured maximum amplitude in the interval before participants spoke (around 1–2 s). Note this interval is quite long but we chose to use this time interval to gain a realistic picture of likely background noise across the recording. Twenty-four tokens from Blackburn (4 %) and 21 from Burnley (3 %) were "noisy", that is, with a SNR less than 10 dB (De Decker 2016: 17). This analysis indicates that 96 % of our data are of adequate quality for acoustic analysis.

We also used the same method to compare our data to: (i) data collected in Lancaster University's noise-attenuated sound booth (n = 478); (ii) data collected in a noise-attenuated sound booth at University College London (UCL; n = 466); (iii) data collected in a quiet room at UCL (n = 455). Datasets (ii) and (iii) are described in Nagamine (2024) and include 12 L1 English speakers from North America producing words containing /l/ and /r/. All data were collected using the same equipment and software described in Section 3.2 (i.e., in AAA with the same microphones and headset). SNRs from all the data are shown in Figure 5.

Differences between the datasets were tested by fitting a linear mixed-effects model to SNR (centred), with fixed effect of dataset and Blackburn as the baseline (Bates et al. 2015). The full model included random intercepts for speaker and word. The overall effect of dataset was tested by model comparison, and differences between datasets were tested via pairwise comparisons in the emmeans package (Lenth 2021). Model comparison showed that there is a significant effect of dataset ($\chi^2(4) = 20.25$, $p < 0.001$). Tukey pairwise comparisons are shown in Table 1 and Figure 6. Blackburn Market has a significantly lower SNR than UCL recording room, and Burnley has a lower SNR than Lancaster sound booth and both UCL recording contexts. There is no significant difference between Blackburn and Burnley Markets. Interestingly, there is also no significant difference between Blackburn

Table 1: Pairwise comparisons for signal to noise ratio in all recording contexts. Slope coefficients with more negative numbers indicate bigger differences in SNR between contexts in dB (centred).

Comparison	$\hat{\beta}$	*df*	*t*	*p* (adjusted)
Blackburn Market – Burnley Market	0.36	68.3	0.42	0.99
Blackburn Market – Lancaster sound booth	−3.99	58.9	−2.64	0.08
Blackburn Market – UCL sound booth	−4.18	59.0	−2.76	0.06
Blackburn Market – UCL room	−4.56	59.2	−3.01	0.03
Burnley Market – Lancaster sound booth	−4.35	58.6	−2.88	0.04
Burnley Market – UCL sound booth	−4.54	58.7	−3.00	0.03
Burnley Market – UCL room	−4.92	59.0	−3.25	0.02
Lancaster sound booth – UCL sound booth	−0.19	51.2	−0.10	0.99
Lancaster sound booth – UCL room	−0.57	51.4	−0.30	0.99
UCL sound booth – UCL room	−0.38	51.4	−0.20	0.99

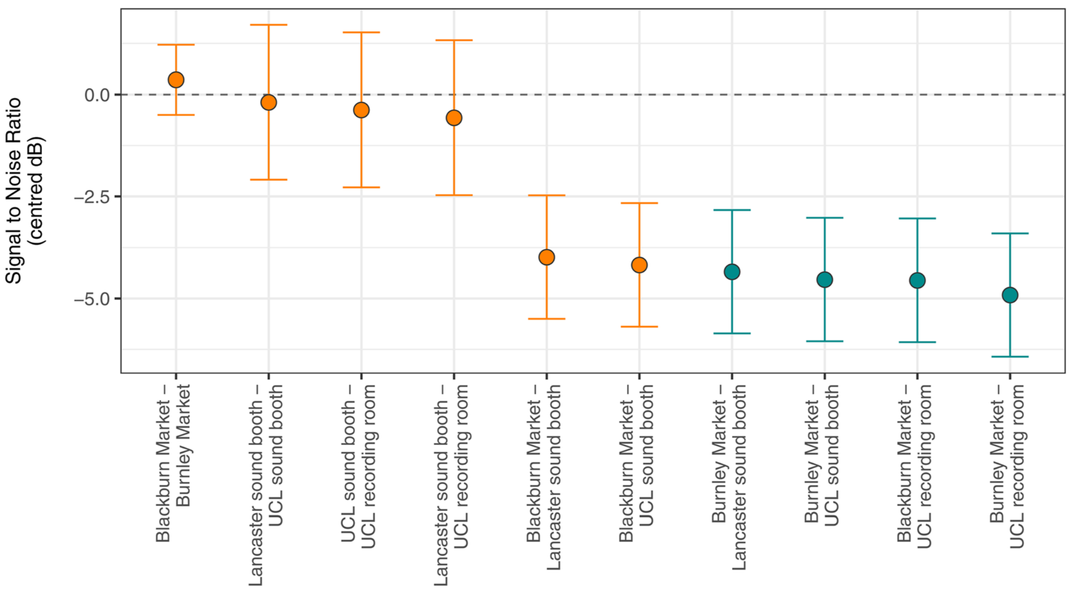

Figure 6: Model output showing pairwise comparisons of signal to noise ratio in different recording contexts. Green comparisons are significant at $p < 0.05$. Error bars show standard error.

Market and Lancaster or UCL sound booth. However, looking at the model output in Figure 6, it is clear that the market settings are similar to one another, the recording booths/rooms are similar to each other, and then the markets differ from the dedicated recording spaces.

3.3.2 Participant demographics

We worked in local market halls in towns which do not have a university. These contexts allowed us to work with people for whom it would not be convenient to travel to a lab on campus. In Blackburn and Burnley, we asked participants their postcode in order to gain an impression of their socioeconomic background. This data allowed us to calculate the Index of Multiple Deprivation (IMD) for their postcode based on English Census data (https://imd-by-postcode.opendatacommunities.org/imd/2019). Of the 64 participants, 49 provided their postcode. Their multiple deprivation decile is plotted in Figure 7. A decile of 1 means that households were in the lowest 10 % of households in England for measures including income, employment, education, health, crime, housing, living environment, and services for children and older people (https://opendatacommunities.org/data/societal-wellbeing/imd/indices). We do not have a comparable dataset with which it would make sense to compare these

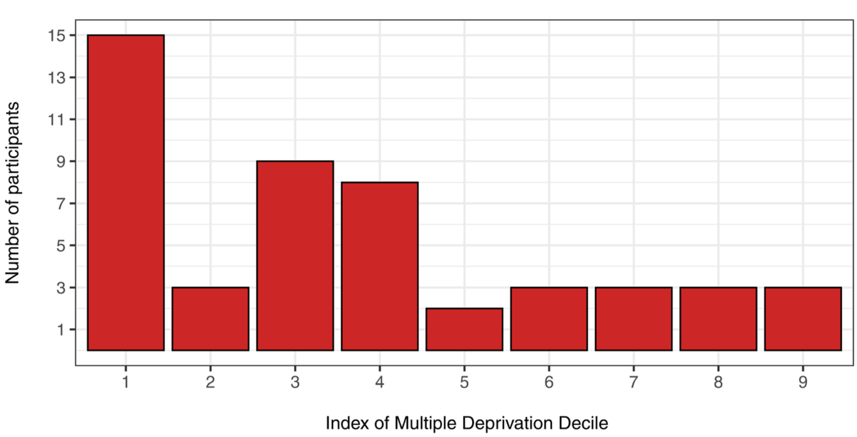

Figure 7: Market ultrasound participants and the Index of Multiple Deprivation decile for their current postcode. Lower numbers = less advantaged, higher numbers = more advantaged.

figures, but it is clear from the data in Figure 7 that we were able to record a large number of participants from areas of multiple deprivation. It is highly unlikely that we would have been able to record this demographic by inviting people to our university campus.

Participants were aged 8–88. We recorded 20 females, 12 males, and no non-binary participants in each location. One participant was transgender. In terms of ethnicity, most of our participants were White British. British South Asians are the largest ethnic minority in East Lancashire. In Burnley, 15 % of residents are British South Asian, and in Blackburn 36 % (Office for National Statistics 2024). In Blackburn we recorded 3 people from the British South Asian community, and 29 White British. At our second event, in Burnley, we made a concerted effort to record more British South Asian participants and included a community member as part of our research team. We recorded 5 British South Asians and 27 White British people. Our sample of 5 British South Asians out of 32 participants (16 %) is reflective of the population in Burnley. We did not record participants from any other ethnic minority. More consideration could be given in the future for how to attract other ethnic minorities.

4 Discussion and conclusion

In this paper we discussed considerations for designing and organizing events that combine public engagement and articulatory phonetic data collection, and then assessed the kinds of data we were able to collect. We found numerous positives from this approach, such as engaging with our local community, a time-effective combination of research and engagement, and data collection from a more socioeconomically diverse population than is typically recorded in our lab. These activities have allowed us to provide paid opportunities for postgraduate students to develop skills in public engagement as well as articulatory phonetics. Also, we now have shared experiences as a lab which have helped us become a more cohesive and collaborative group of researchers.

As discussed in Section 3.1, we carefully considered the kind of research questions that could be answered with this kind of data collection, and the kinds of participants we were likely to recruit in particular venues and locations. Not all projects and research questions will be suited to this kind of data collection. Similarly, as discussed in Section 3.2, it is not feasible to gather large amounts of data per participant in order to maintain a reasonable timeframe for data collection. However, for some research questions this is not necessarily a downside, and we found that our method allowed us to collect a relatively large number of participants in a small space of time. There is more background noise in a market than in a lab setting, but 96 % of our data were above a threshold of background noise considered adequate for accurate acoustic analysis.

Through the discussion and materials provided here, we aim to allow future researchers to consider advantages and disadvantages to this kind of public engagement and articulatory data collection. We hope that our methods can be adapted and improved in order to inspire future work in this area.

Acknowledgments: We would like to thank the anonymous reviewers and editor Yao Yao for their helpful comments on earlier versions of this paper. We received funding from Economic and Social Research Council Festival of Social Sciences, British Academy/Arts and Humanities Research Council Being Human Festival, and Lancaster University Impact Acceleration Account. The collection of data at Lancaster and UCL described in Section 3.3.1 was funded by the Japan Student Services Organization (JASSO), grant ID: 20SD10500601, and a 2022 Research Grant from the Murata Science Foundation, grant ID: M22助人027. Thank you to Blackburn Market and Burnley Market for hosting us and partnering with us for data collection. Thank you to Bronwen Evans for facilitating the data collection at UCL described in Section 3.3.1.

References

Articulate Instruments. 2022. *Articulate Assistant Advanced version 2.20.2*. Edinburgh: Articulate Instruments.

Balch-Tomes, Jonathan & Alan Wrench. 2024. AAA DeepLabCut resources. Github. https://github.com/articulateinstruments/AAA-DeepLabCut-Resources/wiki (accessed 4 July 2024).

Barras, William. 2010. *The sociophonology of rhoticity and r-sandhi in East Lancashire English*. Edinburgh: University of Edinburgh PhD Thesis.
Bates, Douglas, Martin Mächler, Ben Bolker & Steve Walker. 2015. Fitting linear mixed-effects models using lme4. *Journal of Statistical Software* 67(1). 1–48.
Boersma, Paul & David Weenik. 2023. Praat: Doing phonetics by computer, version 6.4.01 [Computer program]. http://www.praat.org/ (accessed 30 November 2023).
Broś, Karolina. 2024. Using social media as a source of analysable material in phonetics and phonology – lenition in Spanish. *Linguistics Vanguard* 9(s4). 349–360.
Conklin, Jenna. 2023. Examining recording quality from two methods of remote data collection in a study of vowel reduction. *Laboratory Phonology* 14(1). 1–28.
Dann, Holly, Sadie Durkacz Ryan & Rob Drummond. 2022. Social meaning in archival interaction: A mixed-methods analysis of variation in rhoticity and past tense be in Oldham. *English Language and Linguistics* 26(4). 861–887.
De Decker, Paul. 2016. An evaluation of noise on LPC-based vowel formant estimates: Implications for sociolinguistic data collection. *Linguistics Vanguard* 2(1). 1–19.
Deliyski, Dimitar D., Heather S. Shaw & Maegan K. Evans. 2005. Adverse effects of environmental noise on acoustic voice quality measurements. *Journal of Voice* 19(1). 15–28.
Freeman, Valerie & Paul De Decker. 2021. Remote sociophonetic data collection: Vowels and nasalization over video conferencing apps. *Journal of the Acoustical Society of America* 149(2). 1211–1223.
Golumbic, Yaela N., Daniela Orr, Ayelet Baram-Tsabari & Barak Fishbain. 2017. Between vision and reality: A study of Scientists' views on citizen science. *Citizen Science: Theory and Practice* 2(1). 1–13.
Heyne, Matthias, Xuan Wang, Donald Derrick, Kieran Dorreen & Kevin Watson. 2020. The articulation of /r/ in New Zealand English. *Journal of the International Phonetic Association* 50(3). 366–388.
Hilton, Nanna Haug. 2021. Stimmen: A citizen science approach to minority language sociolinguistics. *Linguistics Vanguard* 7(s1). 1–15.
Ingrisano, Dennis, Cecyle Perry & Kairsten Jepson. 1998. Environmental noise: A threat to automatic voice analysis. *American Journal of Speech-Language Pathology* 7(1). 91–96.
Kim, Chaeyoon, Sravana Reddy & James Stanford. 2019. Bring on the crowd! Using online audio crowd-sourcing for large-scale New England dialectology and acoustic sociophonetics. *American Speech* 94(2). 151–194.
Kirkham, Sam, Danielle Turton & Adrian Leemann. 2020. A typology of laterals in twelve English dialects. *Journal of the Acoustical Society of America* 147(1). EL72–EL76.
Koreinik, Kadri, Aive Mandel, Maarja-Liisa Pilvik, Kristiina Praakli & Virve-Anneli Vihman. 2024. Outsourcing teenage language: A participatory approach for exploring speech and text messaging. *Linguistics Vanguard* 9(s4). 389–398.
Kostadinova, Viktorija & Matt Hunt Gardner. 2024. Getting "good" data in a pandemic, part 1: Assessing the validity and quality of data collected remotely. *Linguistics Vanguard* 9(s4). 329–334.
Ladefoged, Peter. 1968. *A phonetic study of West African languages: An auditory-instrumental survey*, 2nd edn. Cambridge: Cambridge University Press.
Ladefoged, Peter. 2003. *Phonetic data analysis: An introduction to fieldwork and instrumental techniques*. Oxford: Blackwell.
Lawson, Eleanor & Marie Dokovova. 2023. *AAA introduction manual*. Glasgow: University of Strathclyde.
Leemann, Adrian, Marie-José Kolly & David Britain. 2018. The English dialects app: The creation of a crowdsourced dialect corpus. *Ampersand* 5. 1–17.
Leemann, Adrian, Péter Jeszensky, Melanir Steiner & Jan Messerli. 2020. Linguistic fieldwork in a pandemic: Supervised data collection combining smartphone recordings and videoconferencing. *Linguistics Vanguard* 6(3). 1–16.
Lenth, Russell V. 2021. emmeans: Estimated marginal means, aka least-squares means, version 1.7.3 [R package]. https://CRAN.R-project.org/package=emmeans (accessed 25 November 2024).
Love, Robbie, Claire Dembry, Andrew Hardie, Vaclav Brezina & Tony McEnery. 2022. The Spoken BNC2014: Designing and building a spoken corpus of everyday conversations. *International Journal of Corpus Linguistics*. 319–344. https://doi.org/10.1075/ijcl.22.3.02lov.
Maryn, Youri, Femke Ysenbaert, Andrzej Zarowski & Robby Vanspauwen. 2017. Mobile communication devices, ambient noise, and acoustic voice measures. *Journal of Voice* 31(2). 248.e11–248.e23.
Nagamine, Takayuki. 2024. Formant dynamics in second language speech: Japanese speakers' production of English liquids. *Journal of the Acoustical Society of America* 155(1). 479–495.
Nance, Claire, Maya Dewhurst, Lois Fairclough, Pamela Forster, Sam Kirkham, Takayuki Nagamine, Danielle Turton & Di Wang. 2023. Acoustic and articulatory characteristics of rhoticity in the North-West of England. In Radek Skarnitzl & Jan Volín (eds.), *Proceedings of the 20th International Congress of the Phonetic Sciences*, 3573–3577. Prague: Guarant International.
Office for National Statistics. 2024. Census 2021. England and Wales Census 2021. https://www.lancashire.gov.uk/lancashire-insight/population-and-households/population-and-households-census-2021-articles/census-2021/ (accessed 8 July 2024).
Parikh, Gaurang & Philipos C. Loizou. 2005. The influence of noise on vowel and consonant cues. *Journal of the Acoustical Society of America* 118(6). 3874–3888.
Price, Hazel & Dan McIntyre. 2023. *Communicating linguistics: Language, community and public engagement*. London: Routledge.
Pucher, Michael, Nicola Klingler, Jan Luttenberger & Lorenzo Spreafico. 2020. Accuracy, recording interference, and articulatory quality of headsets for ultrasound recordings. *Speech Communication* 123. 83–97.

Ryan, Sadie Durkacz, Holly Dann & Rob Drummond. 2022. “Really this girl ought to be going to something better”: Rhoticity and social meaning in oral history data. *Language in Society* 52(3). 459–483.

Rymes, Betsy & Andrea Leone. 2014. Citizen sociolinguistics: A new media methodology for understanding language and social life. *Papers in Educational Linguistics* 29(2). 25–43.

Sanker, Chelsea, Sarah Babinski, Roslyn Burns, Marisha Evans, Juhyar Kim, Slater Smith, Natalie Weber & Claire Bowern. 2021. (Don't) try this at home! The effects of recording devices and software on phonetic analysis. *Language* 97(4). 360–382.

Scobbie, James. 2013. *Ultrasound, phonetics, phonology: Articulation for beginners! Workshop at Landelijke Onderzoekschool Taalwetenschap Summer School*. Groningen: University of Groningen.

Scobbie, James M., Eleanor Lawson, Steve Cowen, Joanne Cleland & Alan Wrench. 2011. A common co-ordinate system for mid-sagittal articulatory measurement. *QMU CASL Working Papers* 20. https://eresearch.qmu.ac.uk/handle/20.500.12289/3597 (accessed 25 November 2024).

Sekerina, Irina A., Anna Smirnova Henriques, Aleksandra S. Skorobogatova, Natalia Tyulina, Tatiana V. Kachkovskaia, Svetlana Ruseishvili & Sandra Madureira. 2024. Brazilian Portuguese-Russian (BraPoRus) corpus: Automatic transcription and acoustic quality of elderly speech during the COVID-19 pandemic. *Linguistics Vanguard* 9(s4). 375–388.

Sevilla, Robert Marcelo. 2024. Yiyang Xiang vowel quality: Comparability across two recording media. *Linguistics Vanguard* 9(s4). 335–347.

Smith, Amy, Marie Dokovova, Eleanor Lawson, Anja Kuschmann & Joanne Cleland. 2023. A pilot fieldwork ultrasound study of tongue shape variability in children with and without speech sound disorder. In Radek Skarnitzl & Jan Volín (eds.), *Proceedings of the 20th International Congress of Phonetic Sciences*, 3874–3877. Prague: Guarant International.

Spreafico, Lorenzo, Michael Pucher & Anna Matosova. 2018. UltraFit: A speaker-friendly headset for ultrasound recordings in speech science. In *Interspeech 2018*. Available at: https://doi.org/10.21437/interspeech.2018–995.

Strycharczuk, Patrycja, Manuel López-Ibáñez, Georgina Brown & Adrian Leemann. 2020. General Northern English: Exploring regional variation in the North of England with machine learning. *Frontiers in Artificial Intelligence* 3. 1–18.

Strycharczuk, Patrycja, Susan Lloyd & James Scobbie. 2023. Apparent time change in the articulation of onset rhotics in Southern British English. In Radek Skarnitzl & Jan Volín (eds.), *Proceedings of the 20th International Congress of the Phonetic Sciences*, 3602–3606. Prague: Guarant International.

Svendsen, Bente Ailin. 2018. The dynamics of citizen sociolinguistics. *Journal of Sociolinguistics* 22(2). 137–160.

Tiede, Mark, Suzanne Boyce, Michael Stern, Teja Rebernik & Martijn Wieling. 2024. Production allophones of North American English liquids. In Cécile Fougeron & Pascal Perrier (eds.), *Proceedings of the 13th International Seminar on Speech Production*, 153–157. Autrans, France. https://issp24.sciencesconf.org/resource/page/id/21 (accessed 21 November 2024).

Titze, Ingo. 1995. *Workshop on acoustic voice analysis: Summary statement*. Clearfield, UT: National Center for Voice and Speech.

Turton, Danielle & Robert Lennon. 2023. An acoustic analysis of rhoticity in Lancashire, England. *Journal of Phonetics* 101. 101–280.

Whalen, D. H. & Joyce McDonough. 2015. Taking the laboratory into the field. *Annual Review of Linguistics* 1(1). 395–415.

Wieling, Martijn, Teja Rebernik & Jidde Jakobi. 2023. SPRAAKLAB: A mobile laboratory for collecting speech production data. In Radek Skarnitzl & Jan Volín (eds.), *Proceedings of the 20th International Congress of the Phonetic Sciences*, 2060–2064. Prague: Guarant International.

Zhang, Cong, Kathleen Jepson, Georg Lohfink & Amalia Arvaniti. 2021. Comparing acoustic analyses of speech data collected remotely. *Journal of the Acoustical Society of America* 149(6). 3910–3916.