

LANGUAGE ANALYSIS FOR THE DETERMINATION OF ORIGIN: AN EMPIRICAL STUDY

Paul Foulkes & Kim Wilson

University of York

paul.foulkes@york.ac.uk, k.wilson541@gmail.com

ABSTRACT

Language analysis is now used routinely as part of a battery of tests to assess the legitimacy of asylum claims. The process thus represents a new application of forensic linguistics, and has come to be known as LADO (Language Analysis for the Determination of Origin).

Academic debate on LADO has focused on the value of including native speaker analysts in the process. There are calls for empirical investigation of current methods, and general issues in speech perception and speaker/dialect analysis. It furthermore appears that phonetic analysis plays a relatively small role in LADO reports.

The present study takes a first step towards answering calls to develop empirical work on LADO. It assesses the abilities of different analysts to identify whether speech samples were genuinely produced by speakers of Ghanaian English (GhE). Four groups were tested: LADO professionals, academic phoneticians, phonetics students, and untrained native speakers (NSs) of GhE. Emphasis was placed on phonetic analysis via reference materials outlining phonetic patterns of GhE.

NSs performed best (86% correct), but not significantly better than academics when *unsure* responses were set aside. LADO professionals performed at chance level. They and NSs gave more confident results than academics and students, irrespective of accuracy. False hits were higher than misses for all groups.

The results are not interpreted simplistically as supporting the view that native speakers are necessarily the best placed to conduct LADO analysis, as in several ways the study was not ecologically valid. However, the study reveals the *potential* of native speakers to perform well in the kinds of task central to LADO, as well as the potential of a stronger focus on phonetic analysis.

Keywords: Forensic phonetics, LADO, speaker profiling, Ghanaian English, sociophonetics.

1. INTRODUCTION

Language testing as a component of asylum seeker cases (LADO) was developed in the 1990s as a response to the increase in asylum claims, and the assumption that a rising number of applicants were not from countries they claimed to be fleeing. Around eight government and private laboratories currently conduct LADO. Government labs deal with around 1,000 cases a year, private labs many more [1]. As part of a battery of tests, claimants may be interviewed about their native language(s) and tested on their ability to speak the language(s) they claim to know. Assessing competence in a language is in some cases relatively simple. For example, asylum claimants from Chechnya can be expected to speak fluent Chechen, whereas few bogus claimants know more than a few lexical items. In other cases, however, claimants come from multilingual backgrounds, there may be little or no academic literature on the relevant languages or dialects, and few linguistically-trained analysts are available to assess native speaker competence.

There has been considerable controversy about LADO in academic circles and in the media [4, 8]. Methods of analysis vary across LADO agencies, with no standard testing methods yet established and little documentation available for public scrutiny (but see [1]). There is also dispute over who is most suitable to carry out this work. Debate surrounds the types of professional qualifications required, appropriate training, and the value of including native speaker analysts in the process [1, 5]. Guidelines established for the sake of governments explicitly reject contributions from linguistically-untrained NSs [7]. There are calls for empirical investigation both of currently practiced methods, and more generally into aspects of speech perception and speaker/dialect analysis that underpin identification of linguistic features [5].

Despite a growing literature on LADO, discussion has mostly been based on case studies [1, 2], reviews of previous work on relevant topics

such as dialect identification by non-linguists [5], and polemic. The present study is the first to take an independent empirical approach to key issues in LADO analysis, and involves the cooperation of a number of LADO agencies.

2. RATIONALE

LADO involves the assessment of whether or not an asylum claimant is a genuine native speaker of a particular language or dialect. It is thus reminiscent of a more general task in forensic phonetics, *speaker profiling* [3], albeit with limited questions to be answered. As noted above, there is disagreement on the extent to which native speaker analysts should play a role in LADO cases. It is also apparent from published studies, case reports and personal communication with LADO agencies, that relatively little attention is paid to phonetic or phonological criteria in case analysis. For these reasons a study was designed that (i) could be performed by analysts with different backgrounds, including LADO professionals, and (ii) focused on phonetic criteria. As a first study of its kind [9] the design was intentionally simple. Further studies are planned to extend and refine the approach taken here, and to apply more realistic conditions.

It should be noted that it was not the intention to treat the study as a simple competition between the listener groups. The conditions of the task were not realistic for the LADO professionals: only phonetic reference materials were provided, analysts could not interact with the speakers (which is sometimes possible in LADO cases), and there were no fake samples in which non-Ghanaian speakers attempted to produce GhE.

3. DESIGN

A listening task was designed in which listeners were asked to identify whether or not speech samples were produced by native speakers of Ghanaian English (GhE). GhE was chosen not because of its importance to LADO but because it is relatively well documented [e.g. 6], and native speakers and listeners could be recruited without difficulty. Furthermore, few of the academic or LADO participants had prior experience of work with GhE, hence the task was in this sense deemed to be equally challenging for all groups.

3.1. Speech materials

Seven samples were recorded, five from native speakers of GhE and two foils from Nigerian

English speakers. Recordings were conducted over the telephone, and involved both free speech and retelling of the Cinderella story, the text of which had been sent to the speakers in advance. Recordings were made via a bespoke telephone intercept device fitted to a landline handset operating on the regular UK landline network. The intercept was connected to a Zoom H4 digital recorder. Recordings were made at a sampling rate of 22 kHz and 16 bit depth.

3.2. Listening task

The Cinderella recordings were down sampled to 10 kHz and edited into .wav stimuli of around 2 minutes in duration. These in turn were compiled into a listening task presented to listeners via a web-based interface. The screen presented a label with the speaker/sample number, a play button and scroll bar to control playback of the stimulus, and the question: *Do you believe this person is speaking Ghanaian English?* Five response buttons were provided to capture the respondents' degree of confidence: *highly probable, probable, unsure, unlikely, highly unlikely*. Listeners were instructed to provide an answer to all stimuli and they were asked to provide reasons for their decision using a free text box on the interface screen. They could take as long as they wished to complete the task.

3.3. Listeners

A total of 42 listeners were recruited (Table 1). Native speakers of GhE (NSs) were recruited via contacts at the University of Ghana, Accra. The student group comprised undergraduate and Masters students at the University of York, who had some training in phonetics. The academic group consisted of phoneticians and PhD students, all with at least some experience in forensic phonetics. The LADO professionals were practicing analysts at LADO labs.

Table 1: Listener sample.

group	N
NSs	9
students	15
academics	10
LADO	8
<i>total</i>	42

3.4. Reference materials

A three page reference document was assembled, based on published descriptions of GhE [e.g. 6]. Only information on phonetics and phonology was

included. The level of detail was designed to be accessible to those with elementary training in English phonetics. Examples are shown in Table 2.

Table 2: Examples from training documentation.

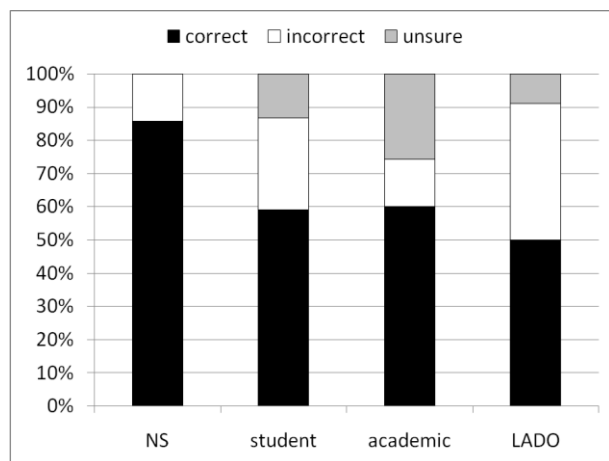
keyword	GhE
NURSE	ɛ(:)
KIT	i > ɪ
commA	a

4. RESULTS

4.1. Overall

Figure 1 summarizes all results. NSs performed best (86% correct). LADO professionals scored 50%. Academics and students were more inclined to use the *unsure* response than other groups: NSs never did so, while 4 of 5 *unsure* responses for the LADO group came from the same analyst. Academics gave 26% *unsure* responses. However, these did not simply equate to ‘don’t know’. Rather, academics indicated that a decision could not be reached with adequate confidence. They typically offered a full explanation of the decision, outlining observed features that matched the training materials, and other features that did not.

Figure 1: Overall results by listener group.



4.2. Decisions reached

Since the LADO task is in essence a forensic one, it should be borne in mind that reaching no firm decision (*unsure* in this experiment) may be the appropriate outcome in cases where materials do not present a consistent or clear picture, and thus no confident conclusion can be reached. It is therefore appropriate to examine results separately for those cases where firm decisions were reached.

Table 3 shows the proportion of correct decisions when *unsure* responses are discarded.

Table 3: Correct results (*unsure* responses removed).

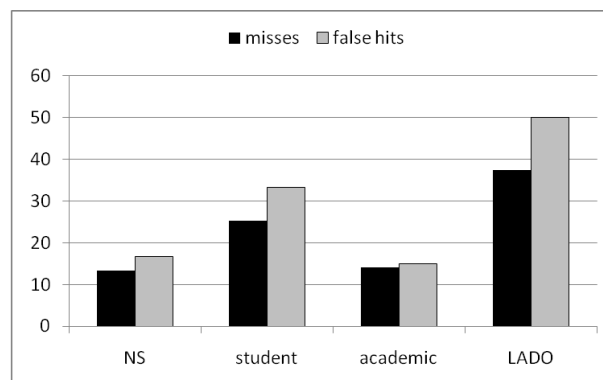
group	% correct
NSs	86
students	68
academics	81
LADO	55

Table 3 shows that correct results for academic phoneticians approached the level achieved by NSs. These data were tested for significance through a logistic mixed effects regression analysis (using the statistical tool R’s LME4 library). Response was included as the dependent variable, listener group as the fixed effect, and listener and sample as random effects. The fixed effect of listener group was returned as significant. Pairwise comparisons were made by fitting the model with different default factors. There was no difference in results for NSs and academics ($p=0.417$). NSs gave significantly more correct responses than LADO professionals ($p=0.0002$) and students ($p=0.007$). Academics gave significantly more correct responses than LADO professionals ($p=0.004$). The difference between academics and students approached significance ($p=0.075$).

4.3. Error patterns

Two types of error were possible: *misses*, i.e. cases where genuine GhE speakers were rejected, and *false hits*, i.e. cases where foils were incorrectly identified as GhE speakers. Figure 2 shows the distribution of error types for each group as a percentage of responses given to genuine GhE speakers (*misses*) or foils (*false hits*). *Unsure* responses were again discarded. For all groups the *false hit* rate was higher than the *miss* rate, most clearly for the LADO professionals.

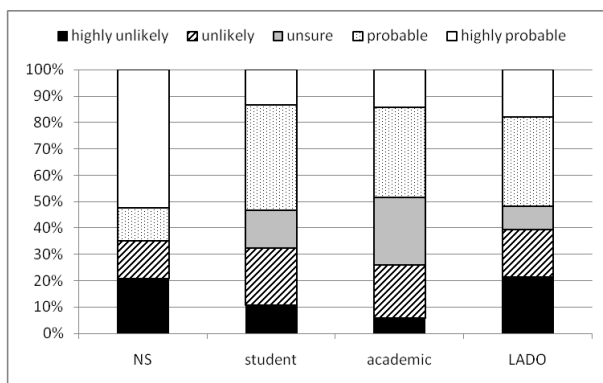
Figure 2: Error types per group as % of responses.



4.4. Confidence

Respondents' confidence ratings were assessed via choices of fixed labels in the listening task (see section 3.2). Results are shown in Figure 3. There was no significant correlation between confidence and accuracy of response. It is noteworthy that responses by the NSs in particular tended to be polarized, with *highly unlikely* or *highly probable* the most frequent responses. LADO professionals also tended to use the extreme labels more frequently than academics or students.

Figure 3: Confidence ratings per group.



5. DISCUSSION

Previous studies show that untrained listeners are not infallible in dialect identification [5]. In this experiment NSs did make mistakes, but they still performed better overall than the other listener groups. This is probably to be expected in this particular task. NSs were also generally more confident in their responses than other groups, with little apparent correlation between accuracy and confidence. In the LADO literature there has been criticism of employing linguistically-naïve NSs as analysts, in particular because they lack the extensive vocabulary to articulate the reasons for their judgments [2]. However, it was apparent from our study that many NSs were indeed able to identify linguistic features that contributed to their decisions, and to identify the foils as Nigerian. Comments such as the following were frequent:

- 'this guy is not speaking GhE, he has Nigerian/Sierra Leonean accent, he pronounces *her* as *haa* and *run* as *ron*'
- 'just isn't GhE. The way she pronounced *Cinderella* and *palace* were give-aways.'

Inspection was made of the NSs' errors. One sample created particular problems, with 44% incorrect responses by NSs. However, academics

fared well with the same sample (10% error). By contrast academics struggled with other samples where NSs did not. These observations lead us to suggest that NSs and academic linguists may pay attention to different cues in identifying GhE, thus an optimal solution might be to combine the efforts of NSs with linguists.

Academics were the most cautious group, articulating their decisions most fully. Correct responses by phoneticians were statistically on a par with those of NSs when *unsure* responses were set aside. Students also performed relatively well. This suggests that detailed consideration of phonetic information is beneficial in LADO cases.

6. CONCLUSION

This study revealed valuable information about the performance of different listener groups in identifying dialect. It marks a first step in empirical investigation of LADO analysis principles. Our initial view is that NSs may have a valuable role to play in the process, provided, of course, that they are proven to have good analytic skills and are supervised by competent linguists.

7. REFERENCES

- [1] Cambier-Langeveld, T. 2010. The role of linguists and native speakers in language analysis for the determination of speaker origin. *International Journal of Speech, Language and the Law* 17, 67-93.
- [2] Eades, D., Fraser, H., Siegel, J., McNamara, T. and Baker, B. 2003. Linguistic identification in the determination of nationality: a preliminary report. *Language Policy* 2, 179-199.
- [3] Ellis, S. 1994. The Yorkshire Ripper enquiry: part 1. *Forensic Linguistics* 1, 197-206.
- [4] Eriksson, A. 2008. Guidelines? What guidelines? Paper presented at IAFFA Conference, Lausanne, Switzerland.
- [5] Fraser, H. 2009. The role of 'educated native speakers' in providing language analysis for the determination of the origin of asylum seekers. *International Journal of Speech, Language and the Law* 16, 113-138.
- [6] Huber, M. 2008. Ghanaian English: phonology. In: Mesthrie, R. (ed), *Varieties of English: Africa, South and Southeast Asia*. Berlin: Mouton de Gruyter, 67-92.
- [7] Language and National Origin Group. 2004. Guidelines for the use of language analysis in relation to questions of national origin in refugee cases. *International Journal of Speech, Language and the Law* 11, 261-266.
- [8] Travis, A. 2003. Language tests to uncover bogus Iraqi asylum seekers. *The Guardian*, 12 March.
- [9] Wilson, K. 2009. Language analysis for the determination of origin: native speakers vs. trained linguists. MSc dissertation, University of York.

Acknowledgments: George Akanlig-Pare, Jamie Garside, Jen Hay, and several LADO agencies.