

Process, tools and agendas in LADO: a rejoinder

Yaron Matras

Abstract

After recapitulating my arguments about methodological flaws in the practices of the company Verified AB in language analysis for the determination of origin (LADO) in asylum procedures, I respond to criticisms of my proposals for an alternative protocol and review the relevance of qualifications, control samples, language variation and scales. I then discuss the interplay of professional competence, ideological bias and understandings of the task and demonstrate the pitfalls of the pre-set framing of the question of the applicants' background.

KEYWORDS LADO, VARIATION, ARABIC, KURDISH, MIGRATION

Affiliation

Aston University, UK
email: y.matras@aston.ac.uk

1. Introduction

The comment by Jim Hoskin, Tina Cambier-Langeveld and Paul Foulkes entitled ‘Improving objectivity, balance and forensic fitness in LAAP: a response to Matras’ (Hoskin, Cambier-Langeveld and Foulkes 2020) critiques my recent article in *IJSLL* (Matras 2018). In that article I addressed inadequacies in the method used by the company Verified AB in reports of language analysis for the determination of origin (LADO) commissioned by the UK Home Office. I wrote the article as an academic who has also had some involvement in casework in recent years around appeals to immigration tribunals in England. I concluded by demonstrating that my critique has been resonating with the courts, which have overturned Home Office decisions that had relied on Verified.

Hoskin et al. focus primarily on my proposals to alleviate methodological flaws and on what they perceive to be the motivation behind those proposals. What they do not do is address the principal content of my article: the risks in the method used by Verified, the major LADO contractor for the UK (and other governments), and the way in which our understanding of linguistic repertoires and language variation can be taken into consideration in LADO analyses in order to mitigate those risks (cf., e.g., McNamara 2019: 191ff.). An opportunity is therefore missed to make progress in forwarding the discussion of methodology in LADO procedures.

In the following I first offer a brief recapitulation of my arguments in the article (Matras 2018). For additional context I then present further insight into the position that the courts in England have taken towards language analysis reports by Verified. I then address the points that Hoskin et al. raise about the disparity among different methodological approaches. Finally, I respond to the suggestion that much of that disparity is driven by incongruent agendas.¹

2. Methods and materials

In my article I assessed a corpus of 50 language analysis reports produced by Verified for the UK Home Office. The reports relate to asylum applicants claiming to be from Syria. Half spoke Arabic and the other half Kurmanji Kurdish. Both are languages that I speak fluently and on which I have carried out research and research supervision over many years. The sample targeted cases where Verified had reached a negative conclusion claiming that the evidence contradicted the applicants’ statement about their place of origin, the Home Office consequently refused the application for asylum, and the applicants appealed to the courts against the Home Office decision. My corpus was thus deliberately selective; the aim was to check the reliability of reports that had led to the refusal of asylum applications.

I should mention here that I have come across a number of reports by Verified that found against the applicant's account and where I agreed with that conclusion, though I had reservations about the method and the interpretation of some of the data. I did not, in the article or elsewhere, claim that all or most reports by Verified arrive at incorrect conclusions or that Verified is responsible for most refusals. (Statistics about the success rate of asylum applications where language analysis is involved are not available).² Rather, my argument was that there were substantial flaws in Verified's method and that therefore there is a real risk that some applications might be unfairly rejected.

My principal points of criticism were these: in none of the reports examined did the authors have knowledge of the language that they were analysing. In the majority of cases, the authors were not trained linguists (i.e., individuals with formal training in descriptive linguistics and language documentation, language change and language variation) but graduates of philological subjects (who had studied, for instance, French or Romanian literature). The native-speaker analysts who supported them generally had no formal training in linguistics. Their input was not verifiable, as it was apparently based on their informal interaction with the authors and did not refer to any accessible control samples or body of notes. There was often incorrect use of published sources and an assumption that absence of evidence (lack of attestation of a form in a certain country) was evidence of absence (proof that presence of that form in the speech sample taken from the applicant necessarily contradicted an origin in that same country). There were incorrect interpretations of individual forms based on lack of awareness of etymology, of distribution conditioned by phonological environment, and of geographical distribution. Such flaws impacted the reliability of the interpretation as to whether a form was consistent or inconsistent with the applicant's stated background. No explanation was provided in regard to the choice of a counter-hypothesis (see below). There was no transparency around the way in which the interpretations of individual examples (some of which were deemed by Verified to be 'consistent' and others 'inconsistent' with the main hypothesis) were used to inform the conclusion, as no qualitative or quantitative assessment threshold was presented. Nor was it clear how the accumulation of impressions translated into a scale representing the outcome. Finally, no allowance was given to the potential complexity of speakers' linguistic repertoires. Variation (referred to invariably in the reports as 'inconsistency') was generally framed as an attempt at deception; no consideration was given to other possibilities, such as the influence of formal or standard speech, stylistic or free variation, special features of border dialects, or a reported history of migration or socialisation in different places.

Hoskin et al. (2020) fail to address most of these points, citing lack of access to the particular cases that I had reviewed,³ though my criticism of the method is

generic and necessarily pertains to all the reports that Verified produces regardless of the conclusions. Instead, they criticise the methodological suggestions that I make toward the end of my article and which aim to alleviate some of the issues that I identify in Verified's reports. Those suggestions can be summarised as follows.

LADO reports should be authored by qualified linguists who are experts in the language that is being analysed. The non-transparent summative result arising from the juxtaposition of two hypotheses should be replaced by an inductive analysis of the data, supported by verifiable control samples. This latter point requires some elaboration: it is agreed that the analyst has a duty to respond to the question put by the client. Almost invariably, that will include the question of whether the linguistic evidence supports the applicant's account of their background or not. If, in the course of the analysis, the analyst comes across evidence that points to an alternative account (i.e., a background other than the one reported by the applicant), then the analyst has a duty to present that account as an alternate hypothesis along with the supporting evidence. That too seems uncontroversial. In the event that the client wishes to have an explicit alternate hypothesis tested, then that too would be included in the analysis (for example, if either the applicant's legal representative or the asylum authorities suspect that the applicant may be from a different background than they claim). In my article I critiqued Verified's practice of framing the analysis as a juxtaposition of two *pre-set* hypotheses: the alternate hypothesis for Arabic speakers was invariably that they were from Egypt and that for Kurmanji invariably that they were from Iraq. None of the Verified reports that I examined indicated how the alternate hypothesis was chosen: no explanation was given as to whether it was requested by the client, the UK Home Office or if it was arrived at inductively by surveying features. I will briefly explain the difference.

If the client asks the expert to test two hypotheses, one that the applicant is from Syria, the other that they are from Egypt, then the expert would state that at the beginning of the assignment. Verified does not do that. Yet it invariably reviews data against just two *pre-set* hypotheses. In that way the presence of the form /mu/ 'not' in an Arabic sample might be deemed to be consistent with Syria; and it would be necessary to point that out if the applicant claimed to be from Syria. On the other hand, the presence of the form /miš/ 'not' would be inconsistent with most regions of Syria (except the south). Now, if the client had asked for a specific alternate hypothesis to be tested, for example that the applicant was from Egypt, then the analysis would point out that /miš/ 'not' is consistent with Egypt. However, if the approach taken was inductive, then /miš/ would have to be considered as consistent not just with Egypt, but also with the dialects of Palestine, Lebanon, Jordan, Libya and more; and so it would not necessarily support

a background just in Egypt. Framing the analyses through pre-set hypotheses therefore has an inevitable effect on the outcome, as it potentially excludes other possibilities. That is what I called in my article a ‘deductive’ approach. I return to the practical issues below in this section, and to the theoretical implications in Section 5.

Hoskin et al. (2020) argue that the necessity to discuss two alternative hypotheses mirrors general Bayesian reasoning in forensic sciences. They cite the Association of Forensic Science Providers (2009: 161), which asserts in connection with criminal proceedings that the expert should address ‘at least one pair of propositions usually one based upon the prosecution issue and one based upon an alternative (defence issue)’. This is referred to as the ‘principle of balance’. A full discussion of the philosophical background of the Bayesian principles is beyond the scope of this rejoinder. As far as I am aware, agencies such as the Dutch government’s language investigation department TOELT and others test whether a single hypothesis based on the applicant’s account (that the applicant is from X) is true or not, rather than offer a particular alternate hypothesis (that the applicant is from Y). Verified’s approach is therefore somewhat particular, and this raises two issues. The first is the quantification of the results. As I point out in the article, citing a concrete example, the way of arriving at a sum of the analyses of individual features in the report, and the overall conclusion derived from those and formulated as a balance between the two hypotheses, is not transparent, and the experiment is therefore not replicable. The second is that Verified fails to state whether its alternative hypothesis/proposition (Egypt, in the case of Arabic) is one that is being introduced by its analysts per se or whether it was the instructing party (the UK Home Office) that is the source of the proposition. Both issues have also been criticised by the courts in England when they overturned Home Office decisions that were based on Verified reports and allowed an appeal in cases of applicants from Syria (see below).

A further point that I make in my article is that, when testing whether the applicant’s account is true or not, the analysis should review features that are diagnostic of regional variation with a view towards **cross-referencing of isoglosses**. This latter point can be illustrated briefly with an example from Arabic: the form /bidd, badd/ for ‘want’ is characteristic of the Levantine dialects of Palestine, Lebanon, Syria and partly Jordan. The negation particle /mu/ in non-verbal predications is found in Syria, Mesopotamia and the northern Arabian Peninsula (Kuwait). The co-occurrence of both forms in a single speech sample thus already points with a considerable degree of certainty towards a socialisation in Syria. Naturally, an analysis would be based on more than just two forms; I cite these for illustration only. The distribution of the forms relied on for comparison can be checked, where available, with reference to recent dialectological studies. For Arabic and

Kurdish, the subject of my analysis, it can also be demonstrated with reference to the dialectological databases available online via the University of Manchester.⁴

This approach can be applied both to the audio recording produced by government contractors such as Verified and to a supplementary recording consisting of two parts: a more extensive sociolinguistic interview tailored to elicit narratives and descriptions (rather than just the listing effect that is typical of elicitations by Verified) and a translation task of around 100 phrases designed to capture diagnostic features more systematically. (Where possible, the translation is from a second language into the applicant's first language; or from Modern Standard Arabic into colloquial dialect). The phrases are extracted from the Manchester database and can be compared directly with control samples. The analysis takes into consideration the possibility of variation and mixture of styles as well as interactional aspects such as hypercorrection and variation conditioned by the turn structure – all issues to which Verified pays no attention at all, even in cases where it is precisely such interactional factors that can reveal an effort by the applicant to emulate another dialect, as I demonstrate in my article (2018: 66–67).

3. Views and verdicts

Hoskin et al. (2020: 268–269) emphasise correctly that the expert's duty is to the court; indeed, in the UK all experts are required to sign a declaration to that effect, which is submitted as part of their expert testimony. The authors also refer to UNHCR guidelines stating that the burden of proof in asylum cases lies with the applicant. According to Hoskin et al., the implication for LADO is that if the linguistic data documented by the expert in an audio recording of around 20 minutes do not contain features that allow the expert to identify unequivocally a particular village or town as the applicant's place of origin, then the applicant has not met that burden of proof. However, such inference is highly problematic: it inevitably makes the outcome dependent on the nature and extent of the data collected by the expert, that expert's aptitude and qualifications, the availability of a verifiable control dataset, and the transparency of the method adopted by the expert to derive an overall conclusion from the interpretation of individual forms – in short, the issues that I address above and in my article. Moreover, Hoskin et al. fail to acknowledge that the legal requirements in relation to the burden of proof may vary (cf. Craig and Zwaan 2019). In England, which is where the reports that I reviewed were commissioned, the courts refer to a 'lower threshold of evidence' by which they generally mean that it is sufficient that the evidence is found to be consistent with the applicant's story. I am aware of 17 recent court determinations (from the years 2017–2021) in cases where Syrians appealed against a Home Office refusal decision that relied on reports by Verified, and I provided counter-expertise. All without exception ended up rejecting Verified's arguments and overturning the Home Office decision. Crucially, the judges accepted my

criticism of Verified's methodology. The determinations are too extensive to cite here in full, but I present a few representative quotations:⁵

I accept the point of Professor Matras that Verified have not mentioned a single item from the appellant's speech that is consistent with Egyptian or Cairo Arabic but not consistent with another Arabic dialect. Even without the expert evidence on language analysis already referred to, I would have had some doubts relating to how Verified had picked a 'hypothesis 2' to work with in this case ... I also accept that criticisms of Professor Matras relating to the Linguist used by Verified in this case are valid. In particular, the Linguist is not said to have any relevant qualifications in the Arabic language and may have no relevant qualifications beyond having attended individual course units in 'Comparative Linguistics'. (First-Tier Tribunal Judge Cruthers, March 2018)

It would be an understatement to say that Professor Matras's report fatally undermines the VERIFIED report relied upon by the Respondent. Professor Matras points out the paucity of qualifications of the 'linguist' and the fact that she is not a qualified linguist at all. He also criticises the methodology used in the VERIFIED report and the quantification of the results obtained ... There are serious deficiencies in the structure of the report and the qualification of the people performing the analysis of the Appellant's speech. (First-Tier Tribunal Judge Foudy, August 2018)

I find that Professor Matras made a number of pointed and well-reasoned criticisms of VERIFIED's report ... I find he has the advantage because he is both a trained linguist and a proficient speaker of the Arabic language ... On balance I find Professor Matras' evidence to be of greater weight ... Professor Matras' training in dialect analysis and his 30 years of research training and expertise is not comparable to 3 native speakers and a linguist with no knowledge of Arabic. It is against that background I also evaluate his critique of VERIFIED's methodology and find that they go to the core and undermine the validity of their conclusions ... Professor Matras' reports are, in stark contrast, transparent with regard to the data used, accessible and his analysis is replicable. As such I must accord his reports greater weight. (First-Tier Tribunal Judge Chowdhury, November 2020).

A recent report by the Independent Chief Inspector of Borders and Immigration in the UK (2020: 86) echoes many of the points made by the tribunals and makes reference to the need for a quality assurance threshold in LADO reports that should include:

the qualifications of the report author team (especially: proven academic training and experience in linguistic analysis in the language that is being analysed), the method of data elicitation and data presentation, and the accessibility of any data samples or sources that the authors rely on for comparison (i.e. the forensic element of the procedure, which requires the availability of a verifiable control sample).

Thus, there is growing concern over the methods and qualifications relied on by Verified among parties involved in reviewing language analysis in UK asylum procedures.

4. Tools and transparency

In their response, Hoskin et al. (2020) misrepresent key aspects of my methodological approach and take issue with others. I use this section to provide clarification.

4.1 The assignment

The authors claim (pp. 259, 264–266) that my initial hypothesis is invariably based on the applicant's story and that I am inclined to attribute inconsistency in dialect forms to a layered socialisation or migration history if the applicant reports such a history. In effect, Hoskin et al. accuse me of refusing to accept that the only explanation for apparent inconsistency is that applicants are not telling the truth. In fact, the method that I endorse is an independent assessment of the data and the **co-referencing of isoglosses**, as demonstrated above. At that point, the expert has a duty to return to the question as formulated by the client in the assignment. If the client's question is whether the data can confirm a socialisation in Syria, then the co-occurrence of the features /bidd/ 'want' and /mu/ 'not', to stay with the above example, would lead me to an affirmative conclusion, assuming the applicant uses those forms correctly and the remaining data are similarly aligned. (By contrast, dialect errors, hypercorrections and direct repetition of forms used by the interviewer contrasting with counterpart forms in other positions in the conversation are common indicators of an effort to emulate another dialect and would lead me to formulate a negative conclusion.) If in addition I encounter the form /kida/ 'such', which is not found in Syria but occurs in Egypt (and elsewhere), then I would need to formulate my conclusion much more carefully and report to the client that the analysis may not offer a clear outcome. If, however, the question put by the client was a different one, namely, whether the data are consistent with a claim to a split socialisation in Syria and in Egypt, then, depending on the balance and distribution of forms in the sample, such admixture of forms in the material may lead me to answer the question in the affirmative.

While the expert's professional and ethical duty is, as we have established, to the court, the expert is also commissioned to address the question put by the client in the assignment. Hoskin et al. conflate the two, insinuating that addressing the client's question when the client represents the appellant is tantamount to neglecting the duty to the court and by implication to the truth. Yet the same conflation can be applied in reverse: Verified also reports to their client (the UK Home Office). And, indeed, the lack of transparency around the choice of counter-hypothesis in their reports raises concerns that this might be based on unsubstantiated suggestions about migration routes and the background of migrants who might not be legally entitled to refugee status. In that respect, addressing two hypotheses rather than just one is not a measure of objectivity but arguably a tool that is built in in order to provide Verified's client with an additional safeguard from its perspective (I return to this point below). My argument about dialect mixture was intended to show that various explanations are possible for what Verified calls 'inconsistency', beyond attempted deception. At the same time, I also discuss a case where I show how hypercorrection and the distribution of forms across turns do indeed point, in my opinion, to an attempt to emulate a Syrian dialect (Matras 2018: 66–67). I believe that my discussion was therefore well balanced.

4.2 Database and analysts

Hoskin et al. (2020: 261) take a critical view of the database on which I draw for comparison samples. They contest that the Arabic speakers who are represented in the database corpus are 'bilingual'; that some of the samples were collected in the UK, that no location is represented by more than two or three speakers, that no further demographics are provided on the Arabic speakers and that for Kurmanji the sample is biased for gender, education and age and so it is not a 'balanced sociolinguistic corpus'. Much of that is factually correct, or at least partly correct. The question is whether these databases offer an advantage over the method that Verified and other agencies rely on and whether they can help safeguard the results from the risk of errors. Let us start by applying the characterisations offered by Hoskin et al. to Verified's native speaker analysts, based on the biographical information provided in the reports that I examined: they are all bilingual individuals, based in Western Europe (usually in Sweden) who have been living there for the past five years and usually much longer. Invariably only one analyst provides input for each of the two hypotheses. No additional control samples are consulted. Moreover, each analyst's input is typically relied on for the entire country from which they originate, that is, the analysts are not chosen to represent the specific town or even the region claimed by the applicant as their place of origin. While no information is provided on gender and age of the analysts, they are all educated, and from the details of their professional background

we can infer that their range of ages is more limited than that represented in the Arabic or even in the Kurdish database. Verified's analysts therefore do not offer any of the qualities that Hoskin et al. identify as potentially lacking in the Manchester database samples.

But the database offers considerably more. First and foremost, the samples are accessible online and are therefore verifiable. That makes the comparison objective and replicable, unlike the informal input of Verified's analysts, which remains internal to their procedure and so escapes scrutiny. The database samples are also comprehensive, albeit in different ways. The Arabic samples each contain around 1,000 tagged and searchable phrases offering opportunities to test the occurrence of a multitude of phonological, morpho-syntactic and lexical forms. Most of the Kurdish phrase datasets are accompanied by free-speech samples available in audio and transcription. The Arabic samples were in most cases recorded from Arabic-English bilinguals. This allows the translation task to be carried out from English as the source language rather than from Arabic, thereby reducing the likelihood that consultants would simply repeat or partly repeat the Arabic phrase as articulated by the interviewer. There is no research evidence to suggest that their knowledge of English impacts in any way on their ability to speak their home dialect. Most of the Kurds are bilingual in the languages of their home region, which is an authentic reflection of the sociolinguistic situation of the majority of Kurds. Samples recorded in the UK constitute a small minority and include for Arabic those recorded by postgraduate students researching their home dialects, and for Kurdish primarily samples belonging to the pilot that preceded the larger-scale data collection in over 120 locations in the Middle East, as the website explains. For Syria, the corpus has Arabic samples representing the three major dialect areas (north, central and south) and Kurdish samples from nineteen different locations representing the major dialect divisions (Afrin, Kobani, Xerbi and Ashiti; cf. Matras 2019). The coverage is thus far more extensive than that offered by Verified's analysts. Moreover, the database samples speak strictly for themselves: they are objective, as they do not purport to make assumptions about other locations. The database thus offers data that are authentic and which, thanks to the elicitation design, offer systematic coverage of key relevant features. By contrast, Verified's analysts are asked to offer their informal and subjective judgment about the speech of locations that they have never visited and from which they do not possess any notes or other documentation. At best they draw on loose impressions retrieved from memory many years after encounters that could have been brief and superficial and are unlikely to have resulted in any systematic coverage of relevant diagnostic features.

In an ideal world we would want to compare a long sample of spontaneous interaction from each applicant with a comprehensive corpus of speech representing variation for gender, age, religion, education, exposure to media, mobility,

multilingualism, and more within their particular claimed place of socialisation; but that is not possible either for Syria or for most regions in the world. It is certainly not possible under the constraints of time and resources imposed by the authorities on their contractors. The opportunity to rely on the Manchester databases must therefore be seen in context as a significant game changer, one that offers major advantages over the faults that the English courts have repeatedly identified in Verified's procedures.

4.3 Variation

The question of variation and multi-layered repertoire is central to the discussion of the extent to which a sample analysis can determine a person's place of socialisation; in my article (Matras 2018) I referred to a number of sources, and I will not reference them again here. Hoskin et al. (2020: 263) note that features that are consistent with Syrian Arabic might also be consistent with another dialect and conclude from this that a second hypothesis is needed. However, a second hypothesis might not be sufficient to equip the expert to address that problem if it is seen as self-contained. For example, the negation marker /mu/ is consistent with Syrian, Iraqi and Kuwaiti Arabic. If the second hypothesis were to be that the applicant is from Iraq, then that would not help us to determine the applicant's likely place of socialisation. But we can do that by **cross-referencing several isoglosses**, as I argue in the article, in an approach that is not necessarily oriented toward a pre-set alternate hypothesis but instead sets out to determine whether the applicant is telling the truth or not, that is, whether the account put forward by the applicant in regard to their socialisation is plausible and supported by the data.

A complication arises when forms might be consistent with different stylistic levels. Thus, Syrian Arabic has /bid̪di/ 'I want' while Iraqi Arabic has /arīd/. But the latter is also Standard Arabic (with variation on the initial vowel). So, when an applicant uses both forms in a recording, a possible and plausible conclusion is that the formality and tension of the interview situation prompts them to resort occasionally to more formal features. Often the speaker is not even aware of such style shifts. Verified notes such cases as 'inconsistencies', but in the reports that I examined it never tries to explain them with reference to the interaction context nor does it explain how it calibrates consistencies in relation to perceived inconsistencies when reaching its overall conclusion. For Hoskin et al., shifts into other styles or potential inter-dialect contact constitute 'secondary influence' (p. 259). They argue that such influence cannot be tested and so it cannot be relied upon. Instead, they maintain, the analysis must seek to establish the 'primary' place of socialisation (p. 266). But this is an unrealistic approach to language. No linguistic theory denies the existence of variation, and no serious and modern theory subscribes to the view of a single-layer 'primary' linguistic socialisation. More-

over, assuming an applicant did, in fact, experience intense immersion in a secondary community having migrated from their place of birth, how would Hoskin et al. propose to isolate their ‘primary’ place of socialisation when analysing their speech? (I challenge the authors to carry out a recording of my English speech to see if they can determine my so-called ‘primary’ place of socialisation). Language is not DNA; it is permeable, and yet permeations can be tested, by taking into consideration insights on sociolinguistic and speech accommodation patterns, power relations in the interview setting and turn-taking structures along with existing documentation on phenomena such as the impact of urban speech on rural communities (cf. Rosenhouse 2013, 2017). Furthermore, we do possess both theory and methods that allow us to study multilingual societies, and incorporating insights from such studies is not only possible but, given the reality of complex linguistic repertoires especially in many of the countries from which refugees originate, it is also a must (cf. Muysken 2019). We have examples of cases where the sociolinguistic reality of complex repertoires has indeed been tested to determine places of socialisation (cf. Blommaert 2009), and we also have methodological models that can guide us to carry out similar analyses (cf. McNamara 2019: 197–211).

4.4 Scales

Finally, Hoskin et al. (2020: 270ff.) contest the absence of a fixed scale of conclusions in the protocol that I propose. Verified opts for phrases such as: ‘The language analysis somewhat suggests that the results obtained more likely than not are consistent with the linguistic community as stated in the hypothesis’. While on the one hand adhering to recommendations that conclusions should not be expressed quantitatively (see Arends and Blommaert 2004), Verified at the same time also uses a numerical scale ranging from +3 to –3. It is often the numerical score that draws the attention of the officials who receive the Verified reports for processing, as can be seen from the content of Home Office decision letters. Scales indicating degree of certainty are not a requirement for objective scientific analysis. Rather, they serve the convenience of the client, who wishes to receive an indication of the report’s reliability and the likelihood that it might withstand a legal challenge by the party that Hoskin et al. metaphorically describe as the ‘defence’. Verified’s scales do not reveal anything about the qualitative or quantitative distribution of features found to be consistent with a hypothesis; in my article (Matras 2018: 62) I demonstrate the mismatch between the sum of conclusions assigned to individual sections in a Verified report and that report’s overall conclusion. Providing a measurable degree of certainty is problematic when there is no statistical corpus on the basis of which the odds of certainty could be calculated. Nor is there an obvious way to decide what weight to give to various structural components even if their quantitative distribution in the relatively small

speech sample is analysed statistically, which is never the case. Thus, while scales may be useful for government clients in order to be able to anticipate the chances of possible appeals, they do not support the accuracy of the academic expertise on which a report is based.

5. Agents and agendas

Towards the end of their comment, Hoskin et al. (2020: 272) raise a set of questions, which I welcome and suggest that they be asked with regard to both the methods employed by Verified and my proposals for an alternative protocol: they ask whether the cause for disharmony among experts might be 1) incompetence, 2) bias or 3) different views of the task. As for the first, I believe that the incorrect interpretation of sources and data in many of Verified's reports speaks for itself: while the input of native speakers can be valuable, it is not a substitute for the supervision of the task through a linguist with training in descriptive linguistics, language change and language variation, and with research expertise in the language that is being analysed.

In regard to bias, I am left intrigued by the authors' indulgence in several excursions: they report on a case in the Netherlands where an accompanying information letter stated that the applicant may have learned Kurmanji as a second language, a statement that Hoskin et al. consider to be speculative (p. 267). It is then reported that the counter-analysis report identified a foreign accent in Kurmanji and concluded, according to Hoskin et al., that the applicant's Kurmanji was 'consistent with the Kurmanji that is spoken in Syria, as well as consistent with her own narrative according to which she acquired Kurmanji as a second language'. Hoskin et al. argue that a speculative statement thus formed the basis for the counter-expert's conclusion. I was able to obtain a copy of the full report via the *IJSL* editor⁶ and discovered that the authors quoted selectively. They neglected to mention that the report author in fact hedged the quoted statement with the words 'as far as one can tell from the material' and then, crucially, stated in the concluding sentence that 'the quality and scope of the material does not allow a more specific conclusion'. Hoskin et al. thus de-contextualised and misrepresented the counter-expertise report.

Similarly opaque is the claim (p. 269) that 'a Dutch court has recently identified a flaw in the general approach taken by Matras'. The authors do not explain what they mean by 'the general approach taken by Matras' nor is that revealed in the short extract that they present from the said court decision, which makes no reference at all to my work. Elsewhere (p. 260), the authors suggest that my critique of Verified was carried out on behalf of De Taalstudio and that my approach is consistent with that described by Verrips (2010). As stated in my article, the material I drew on was obtained from appellants' legal representa-

tives in England. The analysis was part of a research project under my direction that was funded by the Arts and Humanities Research Council. The alternative method that I describe relies on the use of control samples from the Manchester databases, which were developed since 2015, and on the cross-referencing of isoglosses drawing on those control samples. It resembles the one described by Verrips (2010) only inasmuch as both methods aim to offer systematic coverage of relevant linguistic forms under the categories of phonology, morpho-syntax and lexicon and to use that to answer the question that the client has put to the expert. But such approach is common among LADO practitioners.

When it comes to tasks, what Hoskin et al. appear to highlight is in fact the divide between consultants who are contracted by governments and those who are engaged in counter-analysis on behalf of asylum applicants. In suggesting that I take the applicant's story as my point of departure and that I am aligned with those who allegedly adopt speculations as their conclusions, Hoskin et al. are insinuating that my criticism of Verified's methods is driven by an advocacy agenda rather than an academically informed and empirically based approach.⁷ In arguing in favour of concepts such as 'primary' socialisation and brushing aside variation in language as 'secondary', they are dismissing wholesale the appreciation of sociolinguistic reality as well as contemporary linguistic theory that recognises the importance of multiple influences, encounters, mediality and mobility and their effect on individuals' linguistic repertoires.

The issues of bias and understanding of tasks are thus intertwined. Hoskin et al. (2020: 271) admit as much when they decry the fact that agencies are not being paid enough and that governments insist on a quick turnaround of reports. In the context of my critique of Verified reports, such a statement might be interpreted as an admission that Verified's work is error-prone as a result of insufficient dedicated time. This seems to confirm my impression (Matras 2018: 73) that economy of scale may be a factor behind the format of Verified's reports, in particular the framing of the analysis as an experiment to decide between two juxtaposed hypotheses. It is not surprising that governments that are under political pressure to turn away as many migrants as possible, at as low a cost as possible, should seek ways to narrow the chances of applicants to prove entitlement to refugee status (cf. McNamara 2019: 209). Consequently, applicants risk becoming the target of wholesale suspicion. When the analysis is framed as a decision between two hypotheses and the standard pre-set counter-hypothesis is that the applicant is from Egypt, then the format dictates that individual features of speech will be evaluated as consistent or inconsistent with Egyptian Arabic.

Thus, for example, an Arabic-speaking applicant claiming to be from Syria might show a glottal stop in the position representing the historical phoneme /**q*/; Verified would mark this as consistent with the dialect of Cairo, tipping the balance in favour of the counter-hypothesis, while in fact this realisation is

equally consistent with the dialects of Jerusalem, Amman, Beirut and, crucially, Damascus. At the same time, the applicant might pronounce the word for ‘municipal garden’ as /ħadiqa/. In this case the inflectional ending /a/ is conditioned by the phonological environment, as it follows a ‘guttural’ (here: uvular) sound, as well as by the fact that this is a loanword from the Standard language used to designate officially demarcated recreational spaces. In Syrian Arabic, the same inflectional ending otherwise appears as /e/, as in /żibne/ ‘cheese’, whereas in Egyptian Arabic /a/ is retained, thus /gibna/ ‘cheese’. When Verified interprets /a/ in the word for ‘garden’ as consistent with Egyptian Arabic (see Matras 2018: 63) it is therefore introducing a bias (a pre-determined elimination of other possible interpretations) that is conditioned by a particular understanding of the task (the assumption that a pre-set counter-hypothesis needs to be confirmed or refuted). The lack of competence of its analysts and linguist-authors (their lack of familiarity with historical developments in Arabic phonology and the broader pattern of isogloss distribution in the Arabic-speaking world, or their lack of familiarity with Arabic altogether, and their inability to take stylistic variation into account) prevents them from being able to override the process and alleviate the risk of such bias. One way to capture this interplay of institutionalised task-definition and the bias to which it gives rise is to embrace what Gal and Irvine (2019: 167–185) describe as sites of ideologies: a moment in which a view of the world that is partial, dependent on perspective, and so inherently contestable is relied upon to determine boundaries. It is fair to say that Verified’s analysis becomes such a site of ideology, as the framing of the counter-hypothesis and the view of language as monolithic (‘primary’) prescribe a set of boundaries within which the data must be interpreted and preclude other interpretation options.

6. Conclusion

Hoskin et al. (2020) imply that experts who are supposedly working on behalf of asylum applicants or their representatives take a different view of their task and are therefore biased. One could adopt a similar argument and contest that two of the authors have worked as practitioners for agencies that are contracted by governments and that they therefore harbour a bias in favour of curtailing the influx of refugees into Europe. I do not think that either viewpoint is productive. At the heart of the debate lies the question of how linguistic theory and research methodology can best inform LADO practice. That means making optimal use of insights into the dynamics of repertoire permeability among individuals and communities and the impact of multiple encounters, media and mobility on speech; understanding the power relations in the language assessment interview (see already Maryns 2004; De Fina 2019) and applying the tools of conversation analysis wherever these can offer support; making use of the enormous bene-

fits of digital resources for language documentation to share up to date information on an ever-changing dialect landscape; and adopting an approach that ensures objectivity and transparency of comparison samples and data interpretation. Much of this has been argued before, and many of the statements I make to that effect in my article (Matras 2018) are not new, and indeed I reference many other works that have made similar points. My claim to innovation relates to the detailed examination of a corpus of LADO reports and of the outcome of a case where an alternative protocol was applied in a counter-analysis. At the time, I was already able to conclude my article by saying that some of the concerns around LADO practices appeared to be resonating with the courts (Matras 2018: 73). In this rejoinder I have referred to over a dozen court determinations that have since confirmed that this is indeed the case. While the courts cannot be regarded as the ultimate authority on the scientific quality of the linguistic analysis per se, they are entrusted with safeguarding the overall integrity of the asylum procedure, including its reliance on processes for the collection and evaluation of evidence. The position taken by the courts has been that the porousness of Verified's approach is putting asylum applicants at risk and so jeopardising the validity, the fairness and the justice of the procedure for processing asylum applications overall. When Verified's practices are framed as a scientific procedure (cf. Prokofyeva 2019), then this also risks affecting the image and reliability of forensic linguistics as a scientific discipline. A re-evaluation and radical course correction of the method employed by Verified are therefore called for as a matter of urgency.

About the author

Yaron Matras is Honorary Professor at the Aston Institute for Forensic Linguistics, Birmingham, and Guest Researcher at Humboldt University, Berlin.

Notes

1. I am grateful to two anonymous reviewers and to the *IJSLL* editorial team for helpful comments and suggestions on an earlier version.
2. Prokofyeva (2019: 53) claims that the majority of cases that were processed by Verified confirmed asylum seekers' accounts.
3. In agreement with the editors of *IJSLL*, data were presented in the article in such a way that cases and individuals could not be identified. That anonymity was also preserved following publication. I did offer to make available reference numbers of appeals to the courts which I had cited in the article, as I do in this rejoinder.
4. At the time of writing, both databases are still available online and freely accessible. It should be noted that the policy of many UK higher education institutions in regard to online resources is shifting, and there is growing reluctance to continue to maintain online databases without continued

external funding support, which in most cases, once a funded research project has ended, is non-existent.

5. For a longer list of extracts accompanied by the appeal reference numbers, see the tab 'Consultancy' on my personal website: yaronmatras.org
6. I thank Hoskin et al. for passing on the report and the editor for assuming the role of intermediary.
7. Shortly after Hoskin et al.'s (2020) comment appeared in *IJSLL*, it was also published on the website of the company Verified AB, along with photos of the authors and an accompanying commentary. That commentary purports to rely on Hoskin et al. when referring to 'the one-eyedness of professor Matras' and when claiming that 'professor Matras denies the usefulness of working with a native speaker'. The latter claim has no basis in reality; I have often relied in casework on the input of native speakers who were trained linguists at postgraduate and postdoctoral level, and such collaboration formed the basis for the consultancy service MLM-Analysis, which I founded at the University of Manchester in 2017 (for the cited Verified website see <https://www.mynewsdesk.com/se/verified/news/language-analysis-still-under-strain-402897>, accessed on 8 March 2021).

References

- Arends, J. and Blommaert, J. (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(2): 261–266. <https://doi.org/10.1558/sll.2004.11.2.261>
- Association of Forensic Science Providers (2009) Standards for the formulation of evaluative forensic science expert opinion. *Science & Justice* 49(3): 161–164. <https://doi.org/10.1016/j.scijus.2009.07.004>
- Blommaert, J. (2009) Language, asylum, and the national order. *Current Anthropology* 50(4): 415–441. <https://doi.org/10.1086/600131>
- Craig, S. and Zwaan, K. (2019) Legal aspects of LADO from a European perspective: struggling with the burden of proof? In P. Patrick, M. S. Schmid and K. Zwaan (eds) *Language Analysis for the Determination of Origin: Current Perspectives and New Directions* 213–232. Cham: Springer. https://doi.org/10.1007/978-3-319-79003-9_12
- De Fina, A. (2019) The interview as an interactional event In P. Patrick, M. S. Schmid and K. Zwaan (eds) *Language Analysis for the Determination of Origin: Current Perspectives and New Directions* 21–40. Cham: Springer. https://doi.org/10.1007/978-3-319-79003-9_2
- Gal, S. and Irvine, J. T. (2019) *Signs of Difference: Language and Ideology in Social Life*. Cambridge: Cambridge University Press. <https://doi.org/10.1017/9781108649209>
- Hoskin, J., Cambier-Langeveld, T. and Foulkes, P. (2020) Improving objectivity, balance and forensic fitness in LAAP: a response to Matras. *International Journal of Speech, Language and the Law* 26(2), 257–277 <https://doi.org/10.1558/ijssl.39208>
- Independent Chief Inspector of Borders and Immigration (2020) *An inspection of the Home Office's use of language services in the asylum process, May–November 2019* (Open Government License).

- Maryns, K. (2004) Identifying the asylum speaker: reflections on the pitfalls of language analysis in the determination of national origin. *International Journal of Speech, Language and the Law* 11(2): 240–260. <https://doi.org/10.1558/sll.2004.11.2.240>
- Matras, Y. (2018) Duly verified? Language analysis in UK asylum applications of Syrian refugees. *International Journal of Speech, Language and the Law* 25(1), 53–78.
- Matras, Y. (2019) Revisiting Kurdish dialect geography: findings from the Manchester Database. In G. Haig, E. Öpengin and S. Gündoğdu (eds) *Current Issues in Kurdish Linguistics* 225–241. Bamberg: Bamberg University Press.
- McNamara, T. (2019) *Language and Subjectivity*. Cambridge: Cambridge University Press.
- Muysken, P. (2019) Language and origin: the perspective of multilingualism. In P. Patrick, M. S. Schmid and K. Zwaan (eds) *Language Analysis for the Determination of Origin: Current Perspectives and New Directions* 119–130. Cham: Springer. https://doi.org/10.1007/978-3-319-79003-9_7
- Prokofyeva, T. (2019) Linguistic origin identification in focus: theory and practice in LOID. In I. M. Nick (ed.) *Forensic Linguistics: Asylum-Seekers, Refugees and Immigrants* 41–55. Wilmington, DE: Vernon Press.
- Rosenhouse, J. (2013) General and local issues in forensic linguistics: Arabic as a case study. *Comparative Legilinguistics* 15: 53–67. <https://doi.org/10.14746/cl.2013.15.4>
- Rosenhouse, J. (2017) A forensic linguistic problem: asylum seekers' dialect identification difficulties in under-documented adjacent Arabic dialects. *International Journal of Legal Discourse* 2(1): 113–123. <https://doi.org/10.1515/ijld-2017-0005>
- Verrips, M. (2010) Language analysis and contra-expertise in the Dutch asylum procedure. *International Journal of Speech, Language and the Law* 17(2): 279–294. <https://doi.org/10.1558/ijll.v17i2.279>