International Journal of
Language & Law

Hoskin, JLL 13 (2024): 118–147
www.languageandlaw.eu

Evolution and Dialect Perception

— The Case of Language Analysis in the Asylum Procedure

*Jim Hoskin**

Abstract

Little experimental work of direct relevance to Language Analysis in the Asylum Procedure (LAAP) has yet been conducted; neither has theory-building attracted much serious attention. Here I attempt to construct a theory of native speaker competence in dialect perception founded on insights from sociocultural evolution, variationist sociolinguistics, social psychology and antireductionist philosophy. I also augment the limited stock of relevant empirical work with an experimental study involving Yorkshire English as the 'target' dialect and a range of listener groups, from Yorkshire and elsewhere, with and without an educational background in linguistics. Participants in the study (N = 197) were presented with 10, c. 10-second voice samples – four featuring Yorkshire speakers and six featuring non-Yorkshire speakers from the 'linguistic north' of England – and asked 'Is this a Yorkshire accent?' . All participants were native speakers of English. Results of the study suggest that having been born and raised in Yorkshire is the most robust predictor of the ability to accurately perceive Yorkshire speakers. A statistically significant effect on accuracy found for linguistic education is likely attributable to imbalances in the listener sample. I interpret these findings as broadly consistent with the proposed theory of dialect perception, which emphasises bottom-up acuities conferred by the evolution of human sociality rather than – as previously proposed – the enlightening effect of linguistic education. I also discuss the possible consequences of these findings for current approaches to LAAP, considering especially the types of speaker-listeners best suited to perform the linguistic analyses required.

Keywords

language analysis in the asylum procedure, language analysis for the determination of origin, LAAP, LADO, asylum-seekers, dialect perception, dialect recognition, language, evolution

Submitted: 14 January 2024, accepted: 6 May 2024, published online: 18 June 2024

* School of Education, Communication and Language Sciences, Newcastle University ja.hoskin@yahoo.com.
The author would like to thank Professors Ghada Khattab and Paul Foulkes. Research for the article has been kindly supported by an Economic and Social Research Council Postdoctoral Fellowship, awarded by Nine DPT.

1. Introduction

Language Analysis in the Asylum Procedure (herein LAAP, also known by the acronym LADO: Language Analysis for the Determination of Origin) has been used by various governments, especially in Europe, since the mid-1990s. Its purpose is to produce, based on speech data generated in the course of an interview, a linguistic assessment of claims to origin by asylum applicants who are unable or unwilling to prove their identity by conventional documentary means.

The fundamental premise animating LAAP is that false asylum claims are susceptible to discovery through applicants' language production: inauthentic applicants may speak a variety other than that (or those) which might reasonably be associated with their claim to origin, or they may attempt to imitate a variety other than their own. The core question posed in LAAP is thus, in essence, 'What is the likelihood that this person is an authentic speaker of [variety X]?' (but see Patrick (2012) and Matras (2018) for alternative conceptions of the core LAAP question).

At present five LAAP agencies, all located in Europe, are known to be in operation. Verified AB and Sprakab in Sweden are private firms. Verified is contracted by the British, Icelandic, Irish, Swedish, Norwegian, Danish and Finnish governments. As of 2018, Sprakab held contracts with the Austrian, British, Danish, Finnish and Swedish governments (Findahl, 2018). BAMF in Germany, the OCILA in The Netherlands and LINGUA in Switzerland are government bodies.

Verified, Sprakab and the OCILA employ native-speaker non-linguists (NSNLs), in conjunction with trained non-native speaker linguists, to perform language analysis – the so-called 'team approach'. LINGUA accepts as analysts only trained linguists, who may be either native or non-native speakers of the language variety in question. BAMF is known to use some form of automated software, combined with the expertise of human analysts (Deutsche Welle, 2017).

At this juncture I note that my working definition of the term 'native speaker' follows two proposed previously in the literature on LAAP. The first is that of Broeders (2010: 52), for whom a speaker's native identity derives from "… the group of speakers in which he was socialised and learnt to speak his first language."

The second is that of Cambier-Langeveld (2010b: 22):

> […] a native speaker can be defined as a speaker who has first-hand, extensive and continuous experience with the language area and with other speakers of the language and the relevant varieties, starting from an early age.

According to these complementary definitions, 'native speakerhood' is fundamentally a socially acquired, and not a learned or individually elective, category. Thus a speaker of a language, no matter how fluent, cannot be considered a native unless he or she was born and raised in a community in which the language in question is natively spoken and unless he/she has retained contact with some branch of the social organism in which

the language is spoken. The assumption made herein – just as it is, explicitly or implicitly, by LAAP practitioners (cf. Cambier-Langeveld, 2010b; Verified, 2012) – is that native speakerhood comprises a set of language competencies and resulting behaviours which, assuming some degree of enduring embeddedness in the community of socialisation, remain with the individual for life and are thus identifiable via LAAP.

Little experimental work of direct relevance to LAAP has so far been published. Ecologically valid studies conducted hitherto concern the accuracy of NSNL perceptions in LAAP(-like) tasks, involving either closely related linguistic varieties or authentic native vs. imitated speech (e.g. Cambier-Langeveld, 2010a; Foulkes & Wilson, 2011; Hedegard 2015). Most demonstrate the ability of NSNLs to perceive authentic fellow native speakers of their own language variety with high accuracy. Yet, whatever their merits on a purely empirical level, none among this handful of contributions has attempted to develop a theory of native speaker perception that might explain this ability.

This paper thus has four objectives. The first is to develop a theory of NSNL competence in the sorts of dialect perception tasks that are germane to LAAP. The second is to augment the existing stock of empirical work relevant to the field with a new experimental study. The third is to evaluate the consistency of the proposed theory with the results of the new study. The fourth is to assess the implications of these results for LAAP as presently practiced.

With these four objectives in mind, I begin by reviewing material on LAAP of a theoretical nature, contributions of this type constituting by far the majority of published work. I next examine the limited range of more or less pertinent empirical research. Following this, I survey theoretical and experimental material from variationist sociolinguistics, sociocultural evolution, social psychology and antireductionist philosophy. On the basis of the latter overview, and taking my initial cue from the evolutionary account first essayed by Nolan (2012), I then propose two principles of native speaker perception. These in turn inform two hypotheses, the strength of the evidence for which I evaluate through the results of a LAAP-like experimental study, with Yorkshire English as the 'target' variety. In concluding, I sum up the possible consequences of these results for current approaches to LAAP.

My hope is that this dual theoretical and empirical treatment of dialect perception will contribute to a reconsideration of NSNL accuracy in LAAP and provide a persuasive account of its possible source: bottom-up acuities conferred by the evolution of human sociality.

2. LAAP: Theoretical Considerations

The literature on LAAP has, since its inception in the late 1990s, been stalked by discord. Accusations of poor and/or obscure practice are persistently renewed and sporadically

countered. The proper role of sociolinguistic and forensic considerations in posing and addressing the core LAAP question has provoked trenchant disagreement (Hoskin, Cambier-Langeveld & Foulkes, 2020; Matras, 2018, 2021; Patrick, 2012). Most durable have been assertions and counter-assertions as to the competence, or lack thereof, of NSNLs in LAAP.

Some experts (e.g. Eades, 2005; Eades et al., 2003; Fraser, 2009, 2011; Patrick, 2010) argue that NSNLs are deceived by 'folk knowledge' or 'folk views', which are purportedly manifest in an "ideology of homogeneism" (Eades, 2005: 511). Patrick (2010: 77) avers that folk views are "grounded in prescriptive biases (especially educated speakers), and based upon underlying constructs significantly at odds with the facts described by linguistics". The associated ideology of homogeneism, according to Eades (2005), is the commonly-held belief that societies and the people who comprise them are essentially monolingual and monocultural. Resulting misapprehensions about the complexity of language variation are said to render NSNLs incapable of reliably perceiving who is and who is not a genuine speaker of a given dialect.

The perceptions of linguists are apparently free of the spurious linguistic attitudes engendered by the ideology of homogeneism. Linguists alone, it is asserted, know that language, ethnicity and nationality are not always co-determinative categories and that a speaker's place of origin cannot always be decided solely on the basis of the pronunciations or words they use (Eades et al., 2003). For these reasons, it is argued, only linguists can be considered trustworthy judges in LAAP.

In contradistinction, a small but growing body of literature and research by LAAP professionals, forensic phoneticians and sociolinguists takes the view that NSNLs, given supervision and guidance by specialists in forensic phonetics, can be deployed usefully as LAAP analysts (Cambier-Langeveld, 2010a, 2010b, 2012, 2014; Foulkes et al., 2019; Foulkes & Wilson, 2011; IAFPA, 2009; Hedegard, 2015; Hoskin, 2018; Nolan, 2012). The principal theoretical claim in support of this position is that native speakers are especially capable of perceiving subtle distinctions among language varieties which are not captured in segmental phonetic descriptions.

It is, of course, plausible in principle to make out a case that top-down ideologies (e.g. homogeneism/folk views) are decisive in conditioning the perceptions of non-linguists. However, the proposition that NSNLs are impaired by such ideologies in making LAAP judgements is at present without any foundation in ecologically-valid experimental studies. On a theoretical level, critics of NSNLs in LAAP have invoked mainstream ('Labovian' or 'variationist') sociolinguistics (see, e.g., references to Labov in Patrick, 2010). Precisely how the work of variationists demonstrates the unreliability of NSNL judgements in LAAP is never made clear. It appears, though, that the Labovian emphasis on the internally heterogeneous character of language variation (i.e. within as well as among dialects) is taken to suggest that (1) perceiving a speaker's origins is a complex business, which sometimes leads NSNLs to error, and (2) that doing so demands specialist sociolinguistic training. There can be no argument with the former point. As I

argue in section 4, however, the latter can only be upheld based on an incomplete reading of variationist theory.

3. LAAP: Prior Experimental Studies

As previously indicated, experimental research pertinent to LAAP is notable for its scarcity. Fraser (2009) surveys a range of non LAAP-like studies, allowing that "unfortunately, a thorough literature review revealed there is very little research that bears directly on the LADO situation" (2009: 124). Yet she maintains that "judgments about language based on 'folk knowledge', while sometimes accurate, are not reliable enough to be the basis of important decisions" (Fraser, 2009: 118).

In questioning the ecological validity of the works cited and the conclusions drawn by Fraser, Nolan (2012) points out that LAAP demands of analysts a response to a 'one of us?' question (i.e. 'Is the speaker a member of your speech community?') but that many of the experiments Fraser covers involve instead a 'one of them?' task.

Furthermore, the studies described by Fraser more commonly indicate indifferent performance by trained linguists than by NSNLs. It is difficult to understand how such results constitute evidence of the allegedly benighting effects, on NSNLs only, of homogeneism/folk views – for what theory might then account for the poor performance of linguists in the cited results?

Matras (2018) reviews in some detail data he encountered as a 'counter-expert' in a number of real-life LAAP cases initially assessed by NSNL analysts working for Verified, in whose reports he locates multiple inaccuracies. Yet in these cases the true origins of the applicants remain unknown, rendering debatable both Matras's conclusions as to the applicants' identities and, by extension, the theory that top-down factors decisively condition the linguistic perceptions of NSNLs.

Conversely, experimental studies in which the identities of speakers are certain have generally reported that NSNLs are highly accurate in LAAP(-like) tasks. They have even been found to perform more accurately than both specialists in the general linguistics of the 'target' language (Cambier-Langeveld, 2010a; Hedegard, 2015) and expert phoneticians trained to postgraduate level (Foulkes & Wilson, 2011). The sole exception is Muhammad (2021), in which academic phoneticians performed significantly more accurately than NSNLs. If an ideology of homogeneism is indeed operative among NSNLs, it did not in the majority of these studies result in a serious disadvantage in accuracy; on the contrary, it must in fact have functioned to *increase* it. In short, most prior findings call into question the theory that top-down factors serve to diminish the performance of NSNLs in LAAP.

4. Theoretical and Empirical Work Countering the Theory of Homogeneism

4.1. Variationist Sociolinguistics: Structured vs. Unstructured Heterogeneity

As mentioned in section 2, among the theoretical bases for the critique of NSNLs in LAAP is variationist sociolinguistics. However, the extent to which it is valid to draw upon variationism in this connection is contestable.

Essentially, variationism seeks to describe individuals' actual use of language in the context of the speech communities to which they belong. This natural language use is governed by variable forms arranged in "internally structured heterogeneity" (Weinreich et al., 1968). It is axiomatic among variationists that language use varies systematically across parameters beyond those conventionally described in classical dialectological surveys. Differences in the frequency of individual variants (e.g. rhotic vs. non-rhotic pronunciations of certain words, as in the final consonant of southern English vs. Midwestern US *car*) often occur on a spectrum rather than in conformity with clear-cut isoglossic boundaries, and their use may be statistically correlated with the speaker's social class, sex, age group or ethnicity, or by the speech style adopted in a particular context (Foulkes & Hughes, in press). This is what is meant by internally structured heterogeneity, which variationists consider a universal fact of natural language.

Critics of NSNLs in LAAP contend that explicit awareness of this heterogeneity is available solely to (socio)linguists, regardless of whether they are native speakers of a given language variety. Yet this argument ignores two things. First is the value of tacit linguistic knowledge in language perception (cf. below, subsection 4.4.). Second is the fact that speech communities are *particular* social formations whose members ultimately control and are therefore best able to systematically comprehend the unique (albeit socially and temporally dynamic) arrangement of features in the language variety they speak. This is so regardless of heterogeneity among individual speech community members (e.g. their age or sex) and the greater or lesser associated statistical likelihood that they will deploy certain variants in a given context. The point, in sum, is that the available range of variants is generally structured – i.e. constrained by the linguistic norms of a particular speech community – and that native speakers belonging to particular speech communities have a generally inimitable productive and perceptual command of this structure.

Applying this reading of variationist theory to NSNL perceptions in LAAP, it would seem to follow that speakers attempting to imitate a variety other than their own cannot, except in rare individual cases, competently deploy indexical (i.e. either socially- or regionally-marked) variants used by authentic members of the target speech community. Moreover, if one of the characteristics of natural language is structured heterogeneity, and if native speakers of a given variety are those who are fully conversant in it, it follows that they are most capable of perceiving the unstructured heterogeneity characteristic

of language imitation, as well as the subtly different internal heterogeneity typical of closely-related dialects – albeit perhaps not with total accuracy.

Seen from this perspective, the knowledge characteristic of NSNLs is not primarily governed by an ideology of homogeneism. Rather, it consists in part of the ability to distinguish in the act of speaking and by ear, between native and non-native types of variability: structured versus unstructured heterogeneity; authentically native versus non-native speech. The ability of NSNLs to distinguish among members and non-members of a given speech community is crucial for the validity of the existing practice of LAAP – in particular, the 'team approach' used by three of the five known LAAP agencies (cf. section 1). Variationist theory implies that authentic native speakers of a particular language variety will be able to perform competently in the perception experiment detailed in section 6.

4.2. Sociocultural Evolution: The Ultimate Origin of Dialect Perception?

A growing body of theoretical and empirical work in sociocultural evolution suggests that the ability to distinguish authentic from inauthentic speakers of a particular language variety may be ultimately evolutionary in origin (cf. Nolan, 2012 for the first suggestion of the relevance of evolutionary factors to LAAP). This evidence appears to reinforce that adduced previously, from variationist sociolinguistics, that NSNLs in LAAP are eminently capable of perceiving fellow native speakers with high accuracy. The basic propositions of the relevant strand of sociocultural evolutionary thought are that language variation emerged as a signal of group membership, and that the ability to perceive shared patterns of variation assisted individuals in making decisions as to the trustworthiness of potential cooperators (Richerson & Boyd, 2010; Nolan, 2012; Cohen, 2012).

4.2.1. Sociocultural Evolution and Language: Theoretical Accounts

For the majority of human (pre)history, ethnolinguistic groups are thought to have consisted of small groups of relatively closely-related individuals. Dunbar (2011), for example, proposes tribal groups of approximately 1,500 individuals as the domain in which particular language varieties were spoken among prehistoric populations, with clans of c. 150 individuals being the most intensive site of 'cultural social learning', including language.

However, under the impetus of environmental and technological change, such as occurred during the early Holocene or the spread of agriculture in western Eurasia, human groups experienced a generally accelerating pattern of geographical, genetic and cultural reassortment (Foley & Mirazón Lahr, 2011; Richerson & Boyd, 2010). In the resulting conditions of ever-growing social complexity and concomitantly decreased personal

acquaintance, shared packages of language variation may have functioned as a proxy for enduring genetic relatedness among speakers of specific language varieties (Foley & Mirazón Lahr, 2011). It was on this footing that trusting reciprocal relations could most securely be established and maintained in pre-modern societies, easing cooperation among strangers (Richerson & Boyd, 2010).

Distinctive and systematic linguistic variation has the vital advantage over other social markers (such as clothing or hairstyles) that it is especially difficult to fake; therefore, the ability to detect a shared system of linguistic variants, or even single shibboleths, is especially useful in identifying fellow 'ingroup' members (Cohen, 2012). Exercise of this 'gatekeeping' ability tends to hinder infiltration by members of 'outgroups', who are more likely to 'free ride' on ingroup resources or to propagate false information inimical to ingroup interests (Richerson & Boyd, 2010). Groups which developed systematic linguistic variations of their own – and which were able to distinguish variations typical of the ingroup from those typical of outgroups – were thus at a considerable evolutionary fitness advantage relative to those less capable in this domain (Cohen, 2012).

The contribution of individual- and group-level, and of cultural and genetic, factors in the evolution of language variation is disputed (cf. Pinker, 2012). However, a simplified yet credible 'multilevel' synthesis, incorporating all four factors, might be as follows (cf. Haidt, 2012). Greater willingness on the part of individuals to cooperate preferentially with more trustworthy ingroup members, identified by their use of shared linguistic features, resulted in a greater number of mutually-profitable interactions. The relative material prosperity derived from these interactions conferred a fitness dividend – i.e. the production of relatively large number of offspring – to the individuals concerned.

Such individual-level processes were in turn highly consequential for selection at the group level. Transmission among individuals of genes coded for 'groupish' cooperation (i.e. those which conduce to the fitness of the group, even to the short-term cost of the individual), and/or cultural traits which engender groupish behaviour (e.g. markers of group allegiance such as distinctive hairstyles or clothing), increased groups' propensity to cooperate preferentially with fellow ingroup members. These genes co-evolved with the ability, honed by intensive cultural selection, to identify shared linguistic variants as a reliable indicator of individual trustworthiness, which – through a feedback loop with individual- and group-level selection – became more pronounced and prevalent.

According to this account, the tendency of human groups to maintain systematic, and often highly subtle, language variation – and to perceive it as an essential marker of group identity – has deep roots in prehistory. It is not surprising, then, that group-level language variation and the ability of individuals to detect it has endured into the present, even though the attendant advantages are less obvious under the comparatively relaxed selection pressures prevalent in much of the modern world.

4.2.2. Sociocultural Evolution and Language: Empirical Evidence

It is readily apparent that multilevel theories emphasising bottom-up, group-oriented evolutionary processes stand in opposition to the social constructivist theory of homogeneism, which proposes top-down transmission of ideas (in this case formal linguistic education, or the lack of it) to individuals as the fundamental determinant of accuracy in dialect perception. Above I have attempted to demonstrate that the theoretical and empirical bases for homogeneism are dubious. But what of the merits of the empirical evidence for the various evolutionary accounts?

Nettle & Dunbar (1997) conducted a computer simulation of reciprocal exchange between 100 'organisms' of four different types. These were each programmed with a 'dialect', a limited 'memory' and varying permutations of exchange/linguistic strategies. The first type of organism, named COOP, always gives unless its giving has been previously unreciprocated by another organism; the second, POLYGLOT, gives only to 'speakers' of the same dialect and changes its dialect to that of an organism from which it receives gifts; the third, CHEAT, never gives; the fourth, MIMIC, also never gives, but like POLYGLOTs changes its dialect when gifted by another organism. All were modelled as 'non-kin', and CHEATs and MIMICs as free riders.

Results showed that, initially, populations of CHEATS were collectively impoverished, while those of COOPs were collectively wealthy. Unsurprisingly, incursions of CHEATS into COOP populations had a marked collective impoverishing effect. As the memory span of COOPs was increased, CHEATS and MIMICS (the two free-riding populations) were eventually isolated or died out entirely, along with their 'dialects'. Conversely, stable and distinctive dialects emerged along well-established paths of trustful exchange involving COOPs and POLYGLOTs.

These findings are taken by the authors as support for the inference that, in the real world, stable and fecund groups speaking shared language varieties emerge from patterns of trusting reciprocity. This interpretation reinforces the theoretical association made in evolutionary accounts between reciprocal exchange, group-level fitness and the exclusion of free riders via perceptions of linguistic similarity.

As reviewed in section 3, several experimental studies have located highly accurate performance by NSNLs in LAAP(-like) tasks, though none interpret their findings through an evolutionary prism. A study that explicitly attempts to investigate evolutionary theories of accent perception, though not with particular reference to LAAP, is by Goodman et al. (2021). Initial results were drawn from 50 participants from the British Isles who were recorded reading sentences, once in their own dialect and once in imitation of another variety. Other participants then listened to 12 recordings, six of imitators and six of authentic speakers. The task was to identify which were imitated and which were genuine.

All listeners, regardless of local identity, showed a combined 66.7 % probability of correctly rejecting imitators and accepting authentic native speakers of their own local variety. The authors interpret this finding as support for their hypothesis that accents evolved among ancient populations as 'tags' signifying ingroup membership, an indicator of trusting reciprocity. Preliminary results of a much larger study by the same team, involving c. 1,000 participants engaged in the same task, indicate support for local listeners' superior ability to identify imitators of their own variety (70 % to 75 % probability), with non-locals at 55 % to 60 % probability on the same stimuli (J. Goodman, personal communication).

One obvious limitation of the study by Goodman et al., however, concerns the fact that listeners were generally more successful at detecting mimics of accents other than their own: contrary to the study's prediction they were, in other words, more accurate in a 'one of them?' than in a 'one of us?' task. Other qualifications relate to the interpretation of such results. With what degree of security can they be ascribed to temporally-remote evolutionary processes? Can explanations instead be sought in more recent cultural history (e.g. strong public interest in and exposure to regional accent variation in the British Isles)?

The same reservations must be entered in connection with the interpretation of experimental results described in the present paper. Overall, though, the theoretical and experimental work on sociocultural evolution reviewed above may be interpreted as reinforcing the argument, made out previously on sociolinguistic grounds, that native speakers (including NSNLs) are peculiarly able to perceive the linguistic markers defining membership of their own speech community.

4.3. Social Psychology: Further Evidence for Linguistic 'Groupishness' (and Its Automaticity)

Multiple studies conducted in the first two decades of the 21st century have found that respondents from infancy onwards are able to distinguish between same-accent and different-accent speakers, and that they tend to preferentially cooperate with the former (Liberman et al., 2017). Children aged 10 months, for example, are more inclined to either accept gifts from or offer gifts to same-accent than other-accent speakers, (Kinzler et al., 2007). Older monolingual children (five to six years of age) also reveal preferences for same-accent speakers, and strongly associate nationality with accent (Kinzler & De Jesus, 2013). Other research on five- and six-year-olds shows a well-developed perception of dialect gradience (Wagner et al., 2014).

The fact that, even before they can themselves speak, children are able to recognise speakers with the same accent dominant in their personal milieu, as well as evince a greater propensity to cooperate on this basis, suggests that the abilities and prefer-

ences in question may have been mentally hardwired by human evolution. This evidence again points to the conclusion that native speakers, regardless of education, may have an evolutionary-endowed advantage in perceiving same-accent speakers in the LAAP context.

4.4. Antireductionist Philosophy: The Irreducibility of Language Perception

Antireductionist philosophers profess the view that natural phenomena (of which language is one) are unified, hierarchical and emergent in structure; they owe their capacity to function to their interactions as systems, and these interactions give rise in turn to new functions (Polanyi, 1966; Schumacher, 1977). Moreover, according to Polanyi (1966), "we know more than we can tell": many facets of human perception are tacit and thus not susceptible of articulation or reduction to rules.

The antireductionist conception of the irreducible nature of language, and of language perception, essentially parallels conclusions drawn earlier in this article from variationist sociolinguistic theory, sociocultural evolution and social psychology, all of which have been interpreted as countering the theory of homogeneism. How and in what ways this is so can be explained by asking and then answering two questions. First, to what extent might explicit knowledge of a language variety, such as that availed by an education in the reductive discipline of linguistics, be sufficient as a substitute for the tacit perceptions of even an uneducated native speaker? Second, would a linguistic education improve the ability of a native speaker of a particular variety to perceive fellow natives?

Both questions can be answered simply. At the very least, explicit knowledge is different in kind from implicit (or tacit) knowledge. Explicit linguistic education might assist a native speaker in reductively describing the discrete parts of his/her dialect – its phonology, morphosyntax and so on – but it could never act as a substitute for, or fundamentally alter, the holistic perceptions of language-as-identity made available by tacit knowledge.

The insights of antireductionist philosophy, then, are in accord with those gleaned previously from variationist sociolinguistics and sociocultural evolution. Apart perhaps from exceptional individual cases, only native speakers possess the unified, deeply-embedded, tacit perceptions of their own language variety as a system which are necessary to transmit subtle signals of speech community membership, and to fully perceive their social import when they are transmitted by other members.

5. Summary: Two Alternative Principles of Native Speaker Perception and Hypotheses Investigated

In section 2 I outlined the fundamental basis for the critique of NSNL perceptions in the LAAP literature (the theory of homogeneism). This posits the dissemination of hegemonic language ideologies as a crucial factor conditioning the linguistic perceptions of individuals. According to proponents of the theory, the influence of such ideologies can be effectively countered only by advanced education in linguistics, which apparently grants its recipients intellectual independence from the normative pressures exerted on other individuals.

I then examined theoretical accounts and experimental studies relevant to native speaker perception and derived from the fields of sociolinguistics, sociocultural evolution, social psychology and antireductionist philosophy. This examination revealed an alternative view of native speaker perception, the two most important insights of which are set out below in the form of general principles.

(1) Perhaps due to hardwiring by evolutionary selection pressures, native speakers of particular linguistic varieties harbour holistic, implicit, variety-specific knowledge, possession of which makes them especially able to perceive authentic native speakers of the same variety.

(2) Expertise in linguistics, acquired by either native or non-native speakers through formal education, cannot functionally replace native speaker perception of language as an integrated system. However, such expertise may, depending on its nature, assist experts in the description of the elements of language (e.g. phonetics/phonology, morphosyntax).

These two principles inform the hypotheses, formulated below, that are evaluated through the experiments described in section 6.

H1: 'Local' speaker-listeners will show greater accuracy in recognising 'local' voice samples in comparison to speaker-listeners of other ('non-local') varieties.

H2: The predicted differences in accuracy will hold irrespective of whether speaker-listeners are trained in any branch of linguistics.

6. Experiment: Yorkshire English

6.1. Methods and Materials

6.1.1. Task Design

To evaluate the extent of support for the above hypotheses, I conducted an experimental study involving closely-related varieties of English. Speaker-listeners born and raised in Yorkshire were modelled as 'local' listeners, and Yorkshire English as the 'local' (or 'target') dialect. Native speaker-listeners of other English varieties were the 'non-locals', and other northern English varieties were the non-local 'foils'.

Ten short voice samples, each featuring either Yorkshire or non-Yorkshire speakers, were presented to English native speaker-listeners sorted in turn according to (sub)national origin and linguistic education, resulting in the composition of seven listener groups (see below, subsection 6.1.2.). Listeners were then asked to what degree of likelihood they were prepared to associate each voice sample with Yorkshire, on a five-point Likert scale: 'highly likely', 'likely', 'uncertain', 'unlikely' and 'highly unlikely'. A Likert scale was used because it would allow listeners to express their responses with greater nuance than would have been possible had the available range of alternatives been a simple three-way split between 'yes', 'no' and 'uncertain'. However, Likert scale ratings are not considered in the data analysis in subsection 6.2.

The question asked of listeners was, simply: 'Is this a Yorkshire accent?' For Yorkshire born and raised speaker-listeners this is a 'one of us?' question, of the same broad type as that generally addressed by native-speaker analysts in LAAP (cf. section 3).

Before attempting the main task, listeners were first asked, 'Are you a native speaker of English?' Qualtrics' skip logic function prevented progression further into the survey by listeners who answered 'no' to this question. Listeners who had declared themselves native speakers of English then had to answer six further preliminary questions. All except (3) were multiple choice questions; (2) and (6) additionally provided an open-text box in which listeners were to elaborate on their multiple-choice responses.

(1) What is your residential background in the United Kingdom?
(2) Were you born in Yorkshire and/or have you at some stage lived in Yorkshire for more than five years consecutively?
(3) If you answered 'Yes' above and are not originally from the place in Yorkshire where you live now, how long have you lived in your present area of residence? (e.g. 'I was born in London but have lived in Leeds since I was five.')
(4) Which part of Yorkshire are you from/have you mainly lived in?
(5) If you are from the United Kingdom, which part are you originally from?
(6) Have you ever studied dialectology or phonetics at tertiary level or taken a secondary-level course on the accents/dialects of the United Kingdom?

Listeners' answers to these six questions permitted their allocation to one of the seven groups specified below, in subsection 6.1.2. This in turn permitted comparative analysis

of the accuracy in dialect perception of linguistically-trained and -untrained listeners of various degrees of 'localness', and thereby the respective extent of support for H1 and H2. The respective accuracy of the seven listener groups was analysed via a series of binomial logistic mixed effects models (see below, subsection 6.2.1.).

6.1.2. Participants

Participants were recruited through contacts at the University of York and via personal networks, principally on social media, resulting in as representative a sample of educational (tertiary and non-tertiary educated) and vocational profiles as might normally be permitted given the recruitment medium. Data on participant age/sex were not gathered. Ethics permission to conduct the experiment was granted by the Ethics Committee of the Department of Language and Linguistic Science, University of York.

Complete responses were submitted by 198 English native speaker-listeners, who were first allocated to one of four (sub)national categories: Yorks B&R, those born and raised in Yorkshire ('locals'); Yorks Res, those resident in Yorkshire for at least five years but not born in the county; Brit non-Yorks, those resident elsewhere in Britain; and Non-Brit, those resident outside Britain. Each of the first three groups was then stratified into a linguistically-trained and non-trained group. The classification of each individual listener according to linguistic education depended on their response to question (6); cf. subsection 6.1.1. Those assigned to the non-British group were not stratified according to this criterion, since only one listener declared linguistic training in either tertiary-level phonetics or secondary-level training on British dialectology. His/her responses were therefore excluded from the final analysis of results.

The Yorks Res group was composed in order to investigate the 'familiarity effect': the postulate that listeners with a history of secondary residence in a given area tend to show a greater ability, compared to those without such a residential history, to identify speakers originating in their area of secondary residence (Baker et al., 2009, Clopper & Pisoni, 2006). Baker et al. (2009) found that the familiarity effect obtained strongly in listeners who had spent at least five years in their area of secondary residence. A period of at least five years in Yorkshire was therefore determined to qualify listeners for inclusion in the Yorks Res group. Interest in investigating the familiarity effect was motivated by its possible implications for the evolutionarily-informed theory underpinning H1. If a statistically-significant familiarity effect could be located in the analysis of results, the assumption that the attribute of 'localness' reposes only in speaker-listeners born and raised in the 'target' area (here, the county of Yorkshire) would have to be rejected.

Table 1 displays the number of listeners whose responses were included, by education and (sub)national background.

Table 1: Number of listeners, by linguistic education and (sub)national background

	Yorks B&R	Yorks Res	Brit non-Yorks	Non-Brit	Totals
Non-linguists	55	39	45	20	159
Linguists	12	9	17	N/A	38
Totals	67	48	62	20	197

6.1.3. Stimuli

Stimuli (c. 10 seconds in duration, with a range of 9 to 11 seconds) featured excerpted readings of Comma Gets a Cure, downloaded from IDEA: International Dialects of English Archive. Such brief stimuli are appreciably shorter than voice samples encountered in LAAP, which are at least 15 minutes in duration and often up to an hour (Hoskin, 2018; Hubbuch, 2019). However, a duration of c. 10 seconds had emerged from a pilot study and a comparable experiment (cf. Shen & Watt, 2015, in which stimuli of c. 15 seconds were employed) as the optimal length for securing a dataset of sufficient size while making the task manageable for participants.

All stimuli were edited to the requisite length, volume-standardised and noise-reduced using Audacity software. They were then uploaded to Soundcloud, to which they were linked in Qualtrics, the online platform on which listeners took the survey. Stimuli were presented initially in fixed order and then in pseudo-random order. A statistical evaluation of order effects on accuracy was conducted and showed no significant result.

The disadvantages associated with read texts (e.g. risk of shifting to a style closer to the standard, no or virtually no occurrence of regionally marked grammatical or lexical items) were considered reasonable trade-offs for the time and effort spared in composing stimuli from scratch during 2020, when restrictions associated with COVID-19 were in full force. In addition, read texts permit the exercise of greater control over the occurrence of features.

The use of read texts – along with their brevity – nevertheless imposes some limitations on the applicability of experimental results to the practice of LAAP, which relies instead on (semi-)spontaneous speech data derived from an interview with the asylum applicant (these limitations are addressed below, in subsection 6.2.3). A degree of ecological validity was, however, secured by allowing listeners to access stimuli multiple times, as is possible in LAAP.

Of the 10 stimuli, four featured 'target' Yorkshire speakers and six 'foil' speakers from other counties in the dialectological north England. Five of the non-Yorkshire speakers – those from Wigan, Manchester, Salford, Blackpool and Liverpool – originate within the historical (but not present-day administrative) county of Lancashire. The other was of a speaker from Birmingham, a city on the border of north and south whose local accent contains both northern and southern features (Wells, 1982). The Liverpool accent is in many senses quite distinct from those of the rest of historical Lancashire, as well as from those found in the remainder of northern England (Wells, 1982; Watson & Clark 2013).

An explanation is required here of why the above range of stimuli was selected for use in the present experiment. It was thought that the use of Yorkshire stimuli as 'targets' and, for example, stimuli featuring speakers from southern England as 'foils' would make the task too easy for most native speaker-listeners, regardless of their degree of 'localness'. It was also considered, however, that few of the canonical phonetic differences between Yorkshire English and other northern/Midlands varieties (attested in, e.g., Wells, 1982) would be perceptible to most English native-speaking but non-British listeners. In a general sense it was predicted, in line with the hypotheses stated in section 5, that distinctions among the stimuli would be apparent to the listener groups in this experiment on a descending scale of 'localness': Yorks B&R > Yorks Res > Brit non-Yorks > Non-Brit.

Table 2 shows the origins of the speakers in the stimuli and their description in the IDEA database. All stimuli were recorded in the early 2000s. Stimuli names refer to the location in which each speaker had at the time of recording spent the majority of his/her life.

Table 2: Origins of speakers featured in the stimuli and descriptions in the *IDEA* database

Stimulus origin	**Description in *IDEA***
Greetland (Yorkshire)	England 55: female, 91, 1912, white, Greetland and Calderdale (West Yorkshire)
Halifax (Yorkshire)	England 57: male, 52, 1951, white, Leeds and Halifax (West Yorkshire)
Leeds (Yorkshire)	England 81: male, teens, 1989, white, Dewsbury and Leeds (West Yorkshire)
Stainland (Yorkshire)	England 56: male, 48, 1955, white, Stainland (West Yorkshire)
Wigan (non-Yorkshire)	England 15: male, 34, 1966, white, Wigan (Lancashire)
Manchester (non-Yorkshire)	England 71: male, 39, 1968, white, Manchester
Salford (non-Yorkshire)	England 54: female, late 40s, 1950s, white, Salford (Lancashire)
Blackpool (non-Yorkshire)	England 14: female, 28, 1972, white, Blackpool and Liverpool
Birmingham (non-Yorkshire)	England 102: female, 63, 1954, white, Birmingham
Liverpool (non-Yorkshire)	England 44: female, 31, 1972, white, Kirkdale (Liverpool) and Manchester

Note: 'England [#]' refers to database numbering, with standalone numbers referring respectively to speaker age at time of recording and year of birth, and locations to speakers' history of residence.

The geographical origins of speakers in Table 2 are mapped first in Figure 1 (Yorkshire/Lancashire/Midlands shown) and then in Figure 2 (Yorkshire/Lancashire shown).

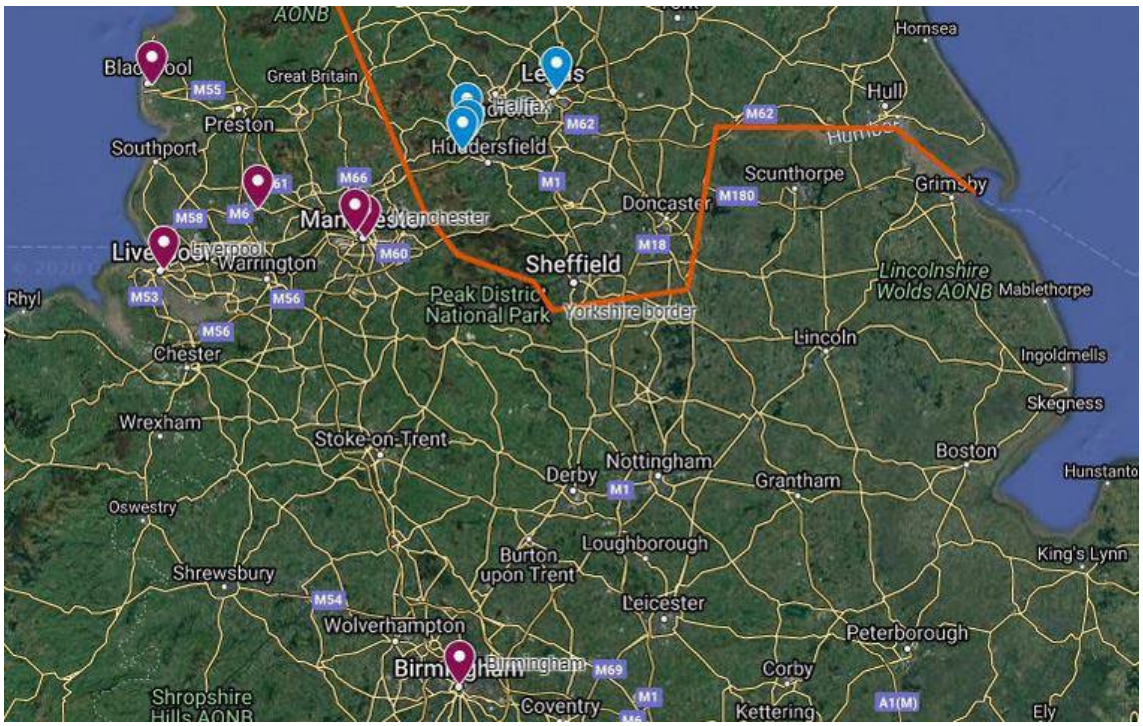

Figure 1: Map of speaker origins; Midlands/Yorkshire/Lancashire partially shown (Yorkshire stimuli in blue, non-Yorkshire in purple; approximate line of historical Yorkshire border in orange)

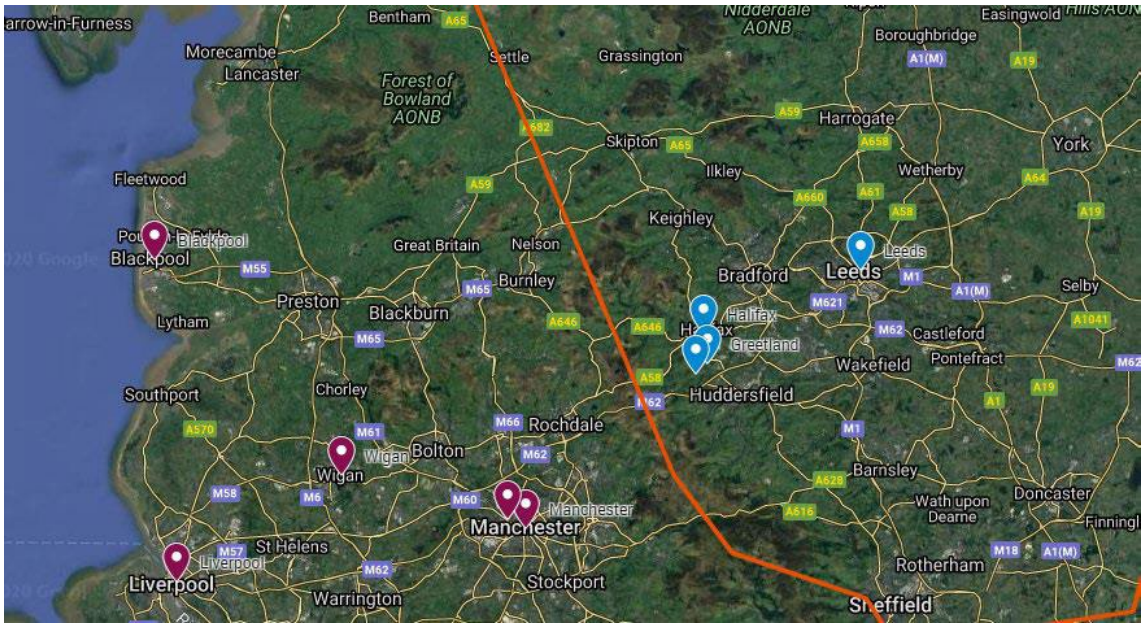

Figure 2: Map of speaker origins; historical counties of Yorkshire/Lancashire partially shown (location of Stainland marked by blue tab southwest of Greetland; Salford marked by purple tab west of Manchester)

The featured excerpt of Comma Gets a Cure was as follows.

> Well, here's a story for you: Sarah Perry was a veterinary nurse who had been working daily at an old zoo in a deserted district of the territory, [so she was very happy to start a new job at a superb private practice in North Square, near the Duke Street Tower].

The section in square brackets was included only in the Wigan stimulus – which, at 11 seconds in duration, was consequently the longest of the stimuli – because it alone featured a realisation of the SQUARE-NURSE merger (as [ə:], in *square*). This vowel merger is reported as typical of parts of northwestern England, including historical Lancashire, but not of most of Yorkshire (Barras et al., 2007; Watson & Clark, 2013). It was thought that its occurrence might serve to distinguish Wigan, the sole non-Yorkshire stimulus of a predominantly 'traditional Northern' character, as non-Yorkshire. This was important because three of the four Yorkshire stimuli were also chiefly of a traditional Northern character and shared many features (but not the SQUARE-NURSE merger as [ə:]) with the Wigan stimulus. However, raw results indicated that inclusion of this feature did not lead to substantially better performance on the Wigan than on any other non-Yorkshire stimulus, with all listener groups performing on it at well below chance level. Below I cite a (non-exhaustive) range of other, potentially diagnostic segmental phonetic features occurring in the stimuli.

The four Yorkshire stimuli featured speakers from Halifax, Leeds, Stainland and Greetland. The first two are urban centres; the others are (semi-)rural areas in Calderdale, between Halifax and Huddersfield. All four stimuli feature typical, though not in all instances exclusively, Yorkshire features. These include – in the final phone of daily and territory – tokens of what Wells (1982) calls 'untensed happY'. This is a feature which is typical of Yorkshire speakers (outside Hull) but which generally occurs also in historical Lancashire, including Manchester but not including Liverpool (Beal, 2010: 18, Baranowski & Turton, 2015: 296). These tokens show intra- and inter-speaker variability but are in all instances monophthongal and with phonetic realisations around [e] or [ɛ]. It was predicted that this feature may be useful to listeners in differentiating Liverpool/ Birmingham speakers, on the one hand, from Yorkshire/Lancashire speakers, on the other.

All four Yorkshire stimuli feature monophthongal FACE in the penultimate vowel of daily, which in Greater Manchester, Liverpool and Birmingham is usually a narrow diphthong, [ei] (Hughes et al., 2012: 117, Baranowski & Turton, 2015: 295). In the rest of historical Lancashire, as in most of Yorkshire, it is predominantly monophthongal (Hughes et al., 2012: 150). Tokens of this vowel were considered fairly strongly diagnostic within the total range of stimuli because only one of the non-Yorkshire stimuli (Wigan) was from Lancashire outside Greater Manchester, as defined on linguistic grounds by Baranowski & Turton (2015).

The four Yorkshire stimuli are exponents of monophthongal GOAT (in *so*). This was predicted to be a strongly diagnostic Yorkshire feature, since it is a diphthong in all the non-Yorkshire stimuli. All the Yorkshire stimuli, except Leeds, are exponents of the GOOSE vowel in *zoo* (and *you*, in the Greetland and Stainland stimuli) as, approximately, [ʊ̈]. This is not present in any of the non-Yorkshire stimuli, where the vowel is monophthongal. Finally, the Halifax stimulus features a traditional northernism (commonly associated by laypeople with Yorkshire, though also present in Lancashire) in the domain of grammar: 2nd person singular *were*. It was thought that the presence of these

three features in particular might be sufficient to mark the stimuli in question as belonging to Yorkshire – readily so for local (Yorks B&R) listeners and less so for others.

There were six non-Yorkshire stimuli. While Manchester and Salford were of a predominantly General Northern English (GNE) character, a variety thought to have emerged from the relatively recent operation of dialect levelling in mid-northern England (Hughes et al., 2012), it was considered that their complete lack of distinctively Yorkshire phonology might be sufficient to mark them out as non-Yorkshire. The Wigan stimulus (with, as discussed above, its potentially diagnostic SQUARE-NURSE vowel merger), was of a traditional Lancashire type.

Some additional exposition is required here of features present in the remaining three non-Yorkshire stimuli. The features of the Blackpool stimulus were (it seemed to me) broadly GNE, with the exception of the speaker's pronunciation of *nurse*, which is approximately [ɛː]. Its presence in the speaker's repertoire may result from her declared period of residence in Liverpool. This pronunciation of *nurse* is common in Liverpool English (Wells, 1982). It was predicted that this feature, in combination with the other 'General Northern but not necessarily Yorkshire' features observed in the stimulus, might be sufficient to identify it as non-Yorkshire.

The Birmingham stimulus featured (in Perry) a characteristically West Midlands diphthongal realisation of happY, [ɜɪ], which is generally not associated with any accent of Yorkshire (Wells, 1982). This was the only stereotypically non-Yorkshire feature in the stimulus but was considered enough to mark it as such.

The Liverpool stimulus contained abundant stereotypical tokens: the stressed vowels in nurse, Sarah and working as [eː], affricated realisations of /t/ (in *veterinary*, *territory*, *deserted* and *district*), initial /d/ (in *deserted*, *daily* and *district*) and /k/ (in *working*). The stimulus was thus thought to be fairly obviously non-Yorkshire.

One other segmental feature, initial /t-/, was relied upon to help listeners distinguish among (some) 'target' and 'foil' stimuli. Foulkes & Hughes (in press) and Turton (pers. comm.) note that an affricated realisation of /t-/ is common in Manchester. Descriptions of similar realisations elsewhere in historical Lancashire are encountered in dialectological accounts – for example, in Bolton (Shorrocks, 1998). Each of the five Lancashire stimuli (Manchester, Wigan, Blackpool, Salford and Liverpool) featured affricated initial /t-/, in territory. By contrast, non-affricated, possibly laminal realisations of initial /t-/ are impressionistically common in traditional Yorkshire speech (Turton, pers. comm.). Though I was unable to locate any published literature on the subject, such realisations of initial /t-/ in territory were indeed present in the Halifax, Greetland and Stainland stimuli.

6.2. Statistical Analysis of Listener Accuracy

6.2.1. Procedure

To compare the accuracy of the various listener groups on both target and foil stimuli, statistical analysis was conducted in R (R Core Team, 2022), by means of a series of binomial logistic mixed effects models (a form of glmer) built with the lme4 package. The three fixed effects were speaker background (two levels: Yorkshire or non-Yorkshire), listener background (four levels: Yorks B&R, Yorks Res, Brit non-Yorks, Non-Brit) and listener education (two levels: linguistic education or no linguistic education). Participant was included as a random effect. Paired interactions between the three fixed effects, at all seven levels, were also incorporated into the full model.
In testing the significance of each variable in the full model, a series of partial models was composed. These were the same as the full model except that in each instance one fixed effect was excluded. Model comparisons were conducted by means of ANOVAs.

Listener responses (i.e. accuracy of response) to the question 'Is this a Yorkshire accent?' constituted the dependent variable. Accuracy was modelled as '0' or '1' according to an incorrect or correct response, regardless of degree of confidence as expressed in Likert-scale ratings.

Outright 'uncertain' responses (i.e. point three on the relevant Likert scale) were excluded from the analysis of accuracy. This decision was made because of the difficulty in deciding whether listeners' selection of this response type constitutes a simple admission of 'don't know' (i.e. lack of competence/familiarity) or caution in identifying stimuli as Yorkshire/non-Yorkshire.

The exclusion of 'uncertain' responses means that the y-axes in Figures 5 to 7 display a higher probability of accuracy than would have resulted from their inclusion. Figures 3 and 4 display raw results for NSNLs and linguists respectively, including 'uncertain' responses, but excluding the full range of Likert-scale ratings. Figure 3 shows that, e.g., Yorks B&R NSNLs achieved c. 75% accuracy on Yorkshire targets but only c. 45% accuracy on non-Yorkshire foils, with c. 10% 'uncertain' responses. This overall pattern is similar in kind though not degree to those of all other groups, non-linguist and linguist, except Non-Brit listeners, who were more accurate on foil than on target stimuli.

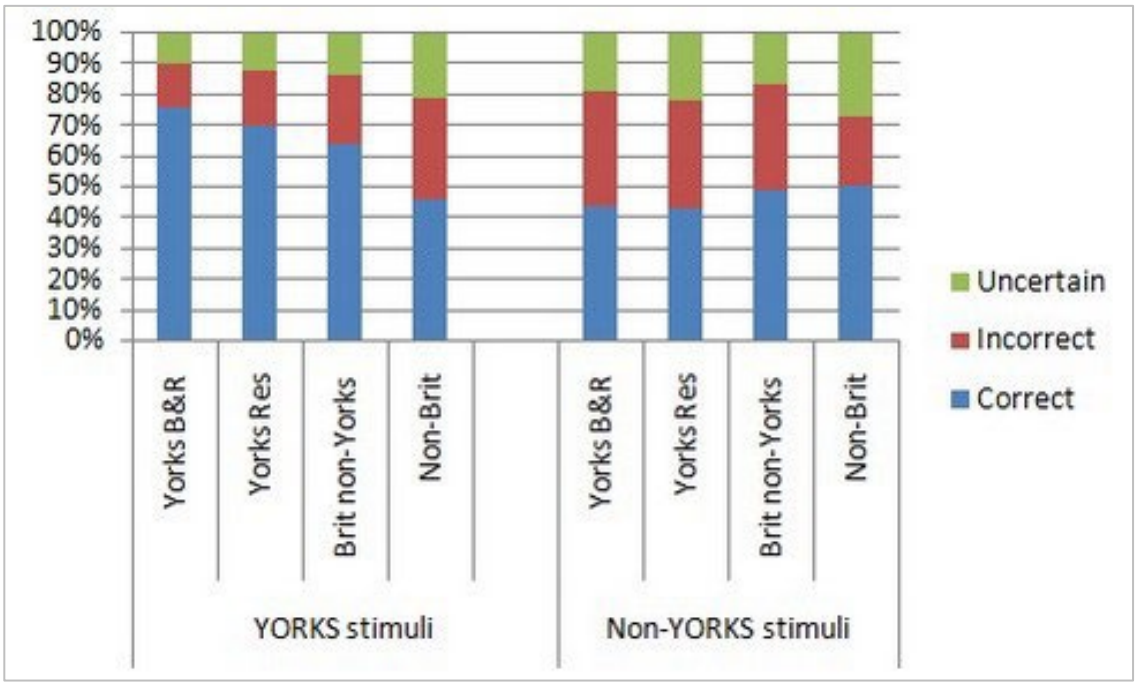

Figure 3: Raw results showing NSNLs' responses to Yorkshire and non-Yorkshire stimuli

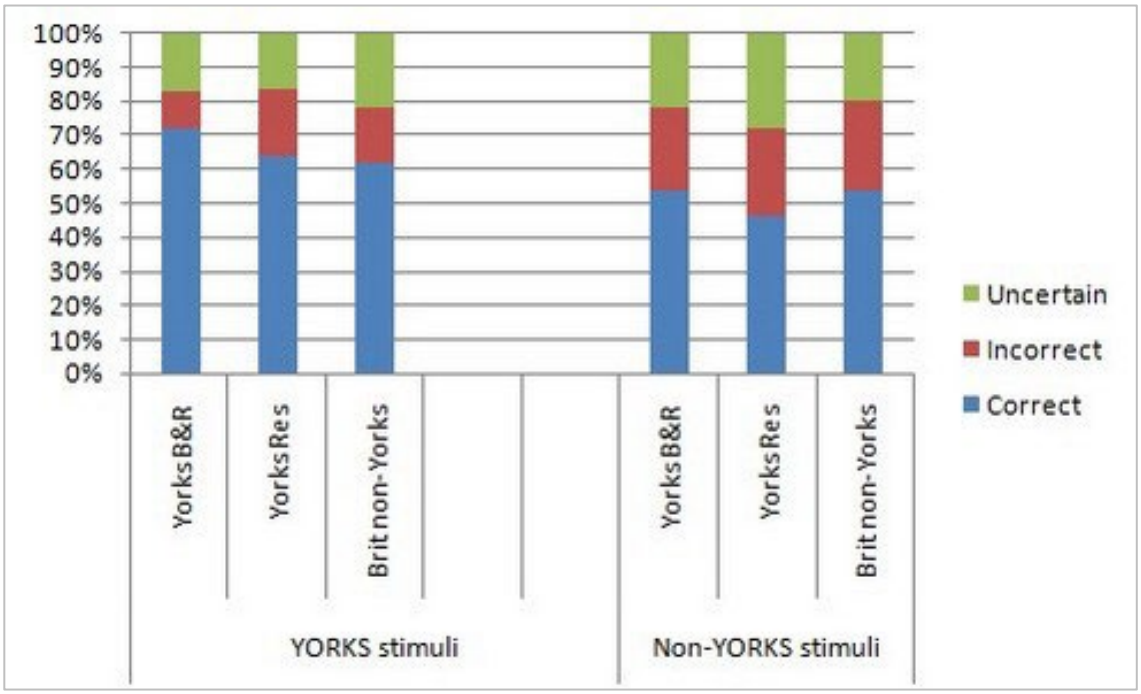

Figure 4: Raw results showing linguists' responses to Yorkshire and non-Yorkshire stimuli

6.2.2. Results: Inferential Statistical Analysis

Table 3 presents the coefficients of the model of best fit and significance values, as determined by glmer, for accuracy of listener response. It also specifies significance values, where present, for the fixed effects.

Table 3: Coefficients and significance values for glmer analysis of accuracy

	Estimate Std.	**Error z**	**value**	**Pr(>|z|)**	
(Intercept)	0.6567	0.1589	4.133	3.58e-05	***
Educationnon-linguist	-0.3064	0.1438	-2.132	0.033040	*
SpeakerYorks	0.6816	0.1985	3.434	0.000595	***
BackgroundNon-British	0.3890	0.2520	1.544	0.122706	
BackgroundYorkshire B+R	-0.1561	0.1623	-0.962	0.336085	
BackgroundYorkshire res	-0.1777	0.1791	-0.992	0.321071	
SpeakerYorks: Background-Non-British	-1.0454	0.3880	-2.695	0.007047	**
SpeakerYorks: BackgroundYorkshire B+R	0.7708	0.2897	-2.661	0.007794	**
SpeakerYorks: BackgroundYorkshire res	0.4963	0.3089	1.607	0.108150	

Signif. codes: 0 '***' 0.001 '**' 0.01 '*' 0.05 '.' 0.1 ' ' 1

Results demonstrate significant effects on accuracy for four independent variables, or pairs of independent variables in interaction. The first is education (i.e. whether the listener declared secondary-level or higher training in British phonetics). The second is speaker origin (i.e. whether the stimulus was of Yorkshire or not). The third and fourth specify interactions between speaker origin and listener background (i.e. the modelled extent of listeners' association with Yorkshire through birth and/or upbringing).
Figures 5, 6 and 7 graphically illustrate the predicted probability of correct response (i.e. accuracy) associated in turn with education, then speaker (i.e. whether the stimulus featured a Yorkshire or non-Yorkshire speaker) and finally the interaction between the background of the speaker in the stimulus and that of the listener.

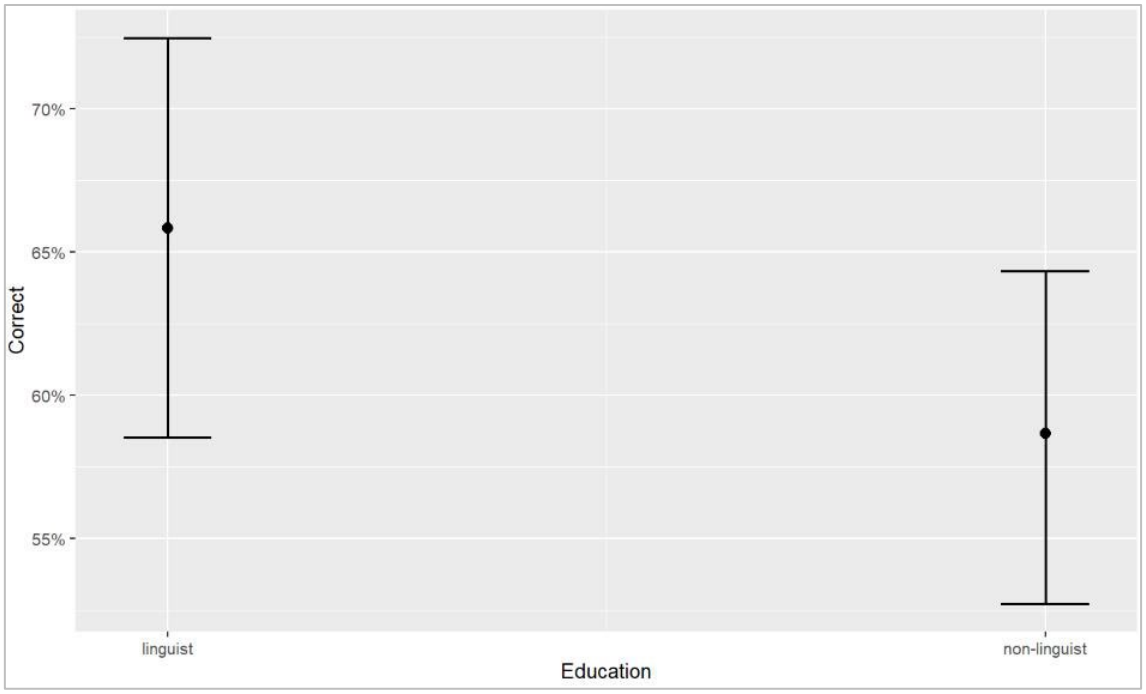

Figure 5: Predicted probability of correct response associated with, respectively, linguistic education and no linguistic education

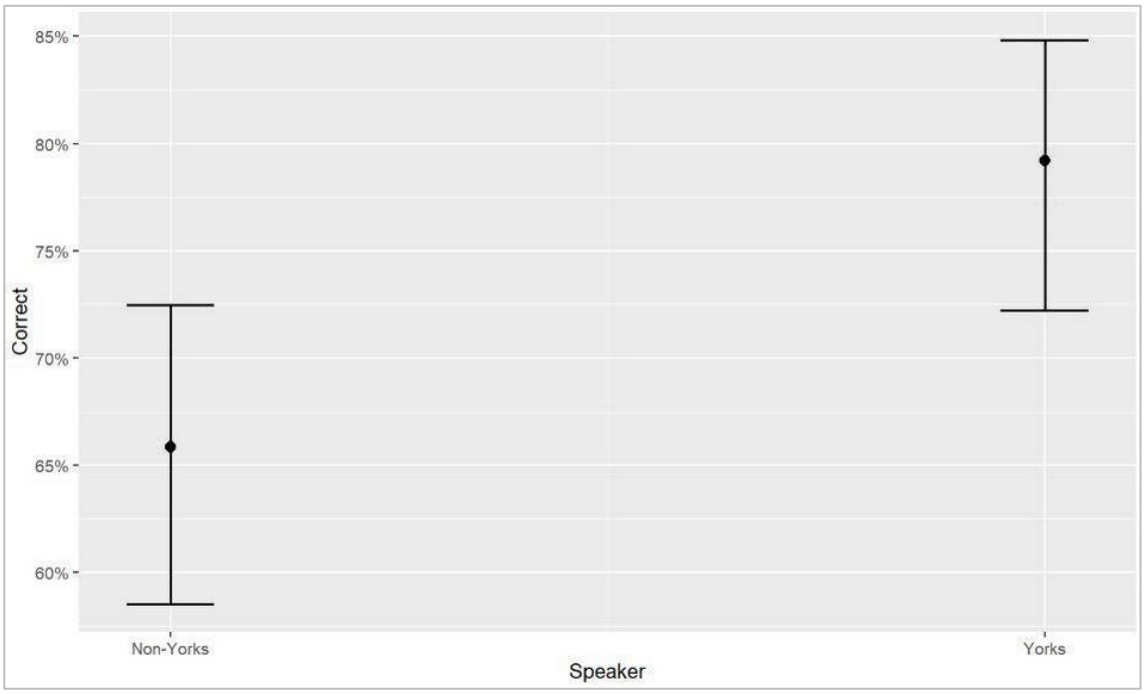

Figure 6: All listeners' predicted probability of correct response associated with, respectively, non-Yorkshire and Yorkshire stimuli

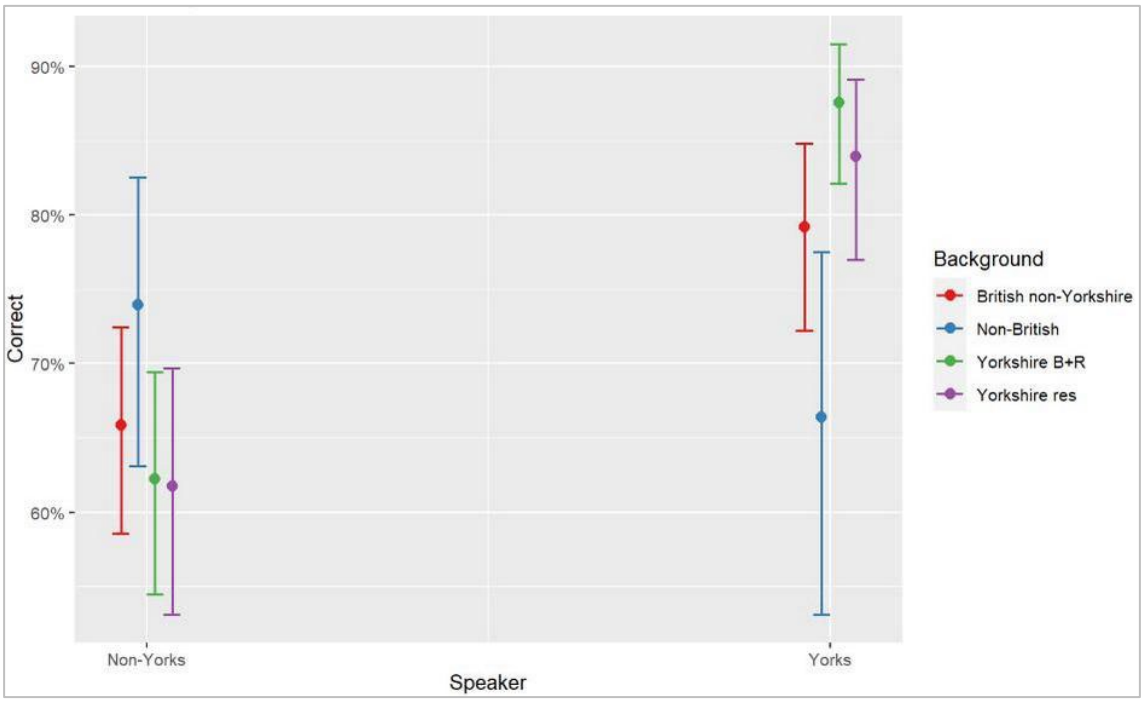

Figure 7: Predicted probability of correct response associated with interaction between listener group and stimulus type

6.2.3. Discussion

The results presented in Table 3 and Figure 7 are somewhat consistent with H1: 'Local' speaker-listeners will show greater accuracy in perceiving 'local' voice samples in comparison to speaker-listeners of other ('non-local') varieties. Yorks B&R were significantly more accurate than other groups, albeit on Yorkshire stimuli only ($p = 0.007$). However, Table 3 and Figure 6 also illustrate a significant effect for accuracy on Yorkshire stimuli for all listeners ($p = < 0.001$). This result must nevertheless be viewed with caution, since in Table 3 the only significant result returned on the interaction between listener and speaker background, other than the positive result involving Yorkshire B&R listeners, was negative (non-Brit, $p = 0.007$). Thus it seems reasonable to infer that the majority of the effect for accuracy on Yorkshire stimuli derived from the response patterns of British groups – Yorks B&R in particular (cf. Figure 7).

The above finding might be interpreted as validation of the familiarity effect, that residency in a particular area makes the local accent easier to perceive (cf. subsection 6.1.2). Indeed, Figure 7 shows a descending probability of accuracy, though again only on Yorkshire stimuli, precisely in line with what would be predicted by the operation of such an effect: Yorks B&R > Yorks Res > Brit non-Yorks > Non-Brit. Nonetheless, this ostensible finding of general support for the familiarity effect must be weighed against the fact that being a Yorks Res listener – a group constructed specifically to explore the validity of the effect – was not a significant predictor of accuracy. The lack of association

between, on the one hand, familiarity with Yorkshire English through residency and, on the other hand, the ability accurately to perceive Yorkshire speakers suggests that the familiarity effect is at best poorly supported by these results.

Figure 7 also shows a close negative association between accuracy and variability in group response patterns to Yorkshire stimuli. The narrowest range was found among the most accurate group (Yorks B&R) and the highest among the least accurate (Non-Brit). On non-Yorkshire stimuli, however, there appears to be a (rougher) positive correspondence between accuracy and variability in response patterns. Here, the Non-Brit group trends highest on both accuracy and variability, while the three British groups trend lower on each. These results together may be interpreted as an indication that the Non-Brit group's response pattern was overwhelmingly random in character (cf. also raw results in Figure 3).

Results therefore appear to be broadly consistent with H1. However, H2 (local listeners will be more accurate in perceiving local speakers irrespective of linguistic education), lacks prima facie support. Table 3 and Figure 5 demonstrate that linguistic training was in fact a significant predictor of accuracy on all stimuli combined ($p = .033$). Yet a likely crucial confound must be mentioned in this connection. This is that the listener groups were unbalanced: there was no non-Brit linguist group (cf. absence from Figure 4), while groups of all three British backgrounds were each divided into a linguist and a non-linguist group. This asymmetry is likely to have skewed inferential statistical results to show greater accuracy on the part of linguists. The extent to which these data are consistent with H2 therefore remains indeterminate, and further investigation of the effect of linguistic education on accuracy in dialect perception is required.

A final finding of interest is that no significant differences in any direction were revealed for intergroup accuracy on non-Yorkshire stimuli. Listeners in general appear to have suffered from a comparative disadvantage in correctly rejecting foil stimuli, and this may have substantially contributed to the high percentage of false acceptances, relative to false rejections, in the above results. This observation, which tends to weaken support for H1, is difficult to account for.

One possible explanation might lie in the brevity of the stimuli (c. 10 seconds). Another might reside in the rather finely-graded distinctions between the accents of some of the Yorkshire and non-Yorkshire speakers featured. A third might have to do with the fact that three of the six 'foils' but only one of the four 'targets' were of a chiefly General Northern English (GNE) complexion, a 'levelled' variety not readily identifiable with any particular location in mid-northern England (cf. subsection 6.1.3). All three factors may, indeed, have acted together: stimuli duration, combined with the overall similarity between Yorkshire and non-Yorkshire stimuli and a 'GNE effect', perhaps made correct acceptance of Yorkshire stimuli easier for participants than correct rejection of (closely related, overall more GNE-influenced) non-Yorkshire stimuli.

However, the brevity of the stimuli does not in itself mean that the results described above are invalid, nor that they are inapplicable to LAAP as practiced at present. There

is no *a priori* reason to assume that an experiment featuring stimuli of longer duration would not yield similar results, though this is naturally a matter for future demonstration. Furthermore, the very fact that stimuli of only 10 seconds in duration exposed meaningful differences in intergroup response patterns suggests that there may be a *gestalt*, pre-rational element in LAAP decision-making (cf. Nolan, 2012). Such might account for the advantage in accuracy shown in the above results for 'local' native speakers over other listener groups, an interpretation that would lend weight to the evolutionarily-oriented account of native speaker perception made out in section 4.

7. Conclusion: Summary of Findings and Possible Implications for the Practice of LAAP

In the introduction to this paper I specified four objectives: (1) to propose, in contradistinction to the theory of homogeneism, a theory of native speaker competence in LAAP-like tasks; (2) to augment, with a new experimental study on Yorkshire English, the fund of ecologically valid empirical work in the field; (3) to evaluate the consistency of the proposed theory in the light of the results of the new study; (4) to consider the possible consequences of my findings for current LAAP practice.

In short, the thesis advanced here is that our remote ancestors may have accrued a hard-wired, evolutionarily-adaptive ability to perceive fellow ingroup members on the basis of linguistic markers of collective identity, enabling them to exclude potential free riders. Such an evolutionarily-conferred acuity might mean that native speakers are able to acquire tacit, holistic knowledge of their own language variety which is of a different kind to the explicit, reductive knowledge availed by linguistic education.

On the basis of this general theory I proposed two hypotheses and evaluated their consistency with reference to the experimental results described in this paper. H1 was that Yorks B&R speaker-listeners would, in comparison to speaker-listeners of other varieties, show greater accuracy in perceiving Yorkshire voice samples. This hypothesis is supported by the results of statistical analysis in subsection 6.2, wherein Yorkshire B&R speaker-listeners, regardless of linguistic education, showed a significant advantage in perceiving Yorkshire stimuli. Interestingly, this result obtained in concert with two others: (1) all listeners combined were significantly more accurate on Yorkshire than non-Yorkshire stimuli; and (2) notwithstanding this, the non-Brit group was significantly less accurate than all other groups on Yorkshire stimuli. This may be taken as a demonstration of the strength of the effect for British (and especially Yorkshire) background on the ability to perceive Yorkshire stimuli.

Somewhat vitiating support for H1, however, was the fact that no listener group showed a statistically significant advantage in rejecting non-Yorkshire foil stimuli. As

discussed in subsection 6.3.3, this result might be attributable to the brevity of the stimuli, the lack of distinctively non-Yorkshire features in three of the six foils but in only one of the four Yorkshire targets, or some combination of these factors.

The second hypothesis evaluated (H2) was that the predicted difference in accuracy between locals and non-locals would obtain regardless of whether speaker-listeners had been trained in any branch of linguistics. This hypothesis was partially supported: Yorks B&R speaker-listeners significantly outperformed other groups, on Yorkshire stimuli only, irrespective of linguistic education. Once more, though, I emphasise a complication: the statistically significant effect of linguistic education on accuracy in identification of all stimuli, a finding I have interpreted as an artifact of design anomalies in the present experiment (cf. subsection 6.2.3).

For current approaches to LAAP, three tentative inferences follow from the above results. First, LAAP agencies appear justified in employing NSNL analysts provided that their linguistic background (through upbringing, not subsequent acculturation) closely matches that of the applicant. This inference is generally consistent with those drawn from previous experiments, involving a range of languages, on NSNL judgements in LAAP-like tasks.

Second, agencies (such as LINGUA) which employ non-native linguists as LAAP analysts may likewise be best served by stipulating that all analysts be native speakers whose dialect matches that claimed by the asylum applicant.

Third, the fact that no group significantly outperformed any other in correctly rejecting non-Yorkshire stimuli suggests that the input of local speaker-listeners is vital for securing the maximum number of correct acceptances, with no significant reduction in the number of correct rejections.

Nevertheless, the caveats entered in subsection 6.2.3 and again in this section remain – especially as to the uncertain effect of linguistic education on listener accuracy, as well as the poor general performance of listeners in rejecting non-local stimuli (this having perhaps had to do with stimuli duration and/or the similarity of target to foil stimuli; cf. also subsection 6.2.3). These factors, like the hypotheses investigated and the practical inferences made above, are worthy of further empirical examination elsewhere.

References

Baker, Wendy, Eddington, David & Nay, Lyndsey (2009). Dialect identification: the effects of region of origin and amount of experience. *American Speech*, *84*(1), 48–71. DOI: 10.1215/00031283-2009-004.

Baranowski, Maciej & Turton, Danielle (2015). Manchester English. In Hickey (Ed.), *Researching Northern Englishes* (pp. 293–316). Amsterdam & Philadelphia: John Benjamins.

Barras, Will, Honeybone, Patrick & Trousdale, Graeme (2007). Is word-frequency a factor in phonological demergers? The sociophonology of SQUARE and NURSE in Lancashire English. *Paper presented at ICLCE2*. Toulouse, France.

Beal, Joan C. (2010). *An Introduction to Regional Englishes: Dialect Variation in England*. Edinburgh: University Press.

Broeders, Ton (2010). Decision making in LADO – A view from the forensic arena. In Zwaan, Muysken & Verrips (Eds.), *Language and Origin: The Role of Language in European Asylum Procedures: Linguistic and Legal Perspectives* (pp. 51–60). Nijmegen: Wolf Legal Publishers.

Cambier-Langeveld, Tina (2010a). 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*(1), 67–93. DOI: 10.1558/ijsll.v17i1.67.

Cambier-Langeveld, Tina (2010b). The validity of language analysis in the Netherlands. In Zwaan, Muysken & Verrips (Eds.), *Language and Origin: The Role of Language in European Asylum Procedures: A Linguistic and Legal Survey* (pp. 21–33). Nijmegen: Wolf Legal Publishers.

Cambier-Langeveld, Tina (2012). Clarification of the issues in language analysis: A rejoinder to Fraser and Verrips. *International Journal of Speech, Language and the Law, 19*(1), 95–108. DOI: 10.1558/ijsll.v19i1.95276.

Cambier-Langeveld, Tina (2014). State-of-the-art in language analysis: A response to the chapter on LADO in the Oxford Handbook of Language and Law. *International Journal of Speech, Language and the Law, 21*(2), 371–381. DOI: 10.1558/ijsll. v21i2.371.

Clopper, Cynthia G. & Pisoni, David B. (2006). Effects of region of origin and geographic mobility on perceptual dialect categorization. *Language Variation and Change, 18*(2), 193–221. DOI: 10.1017/S0954394506060091.

Cohen, Emma (2012). The evolution of tag-based cooperation in humans: The case for accent. *Current Anthropology, 53*(5), 588–616. DOI: 10.1086/667654.

Deutsche Welle (2017). *Speech Software to Test Refugees' Claims of Origin*. Available at dw.com/en/automatic-speech-analysis-software-used-to-verify-refugees-dialects/a-37980819 (accessed 12 May 2023).

Dunbar, Robin Ian MacDonald (2011). Constraints on the evolution of social institutions and their implications for information flow. *Journal of Institutional Economics, 7*(3), 345–371. DOI: 10.1017/S1744137410000366.

Eades, Diana (2005). Applied linguistics and language analysis in asylum seeker cases. *Applied Linguistics, 26*(4), 503–526. DOI: 10.1093/applin/ami021.

Eades, Diana, Fraser, Helen, Siegel, Jeff, McNamara, Tim & Baker, Brett (2003). Linguistic identification in the determination of nationality: A preliminary report. *Language Policy, 2*(2), 179–199. 10.1023/A:1024640612273.

Findahl, John (2018). Collecting contemporary knowledge on Gorani spoken east of Mosul. In Nick (Ed.), *Forensic Linguistics: Asylum-seekers, Refugees and Immigrants* (pp. 57–74). Malaga: Vernon Press.

Foley, Robert & Mirazón Lahr, Marta (2011). The evolution of the diversity of cultures. *Philosophical Transactions of the Royal Society: Biological Sciences, 366*(1567), 1080–1089. DOI: 10.1098/rstb.2010.0370.

Foulkes, Paul, French, Peter & Wilson, Kim (2019). LADO as forensic speaker profiling. In Patrick, Schmid & Zwaan (Eds.), *Language Analysis for the Determination of Origin: Current Perspectives and New Directions* (pp. 71–89). Cham: Springer. DOI: 10.1007/978-3-319-79003-9_5.

Foulkes, Paul & Hughes, Vince (in press). Dialectological and sociolinguistic foundations of forensic speaker comparison. To appear in Nolan, McDougall & Hudson (Eds.), *The Oxford Handbook of Forensic Phonetics*. Oxford: University Press.

Foulkes, French & Wilson (2019). LADO as forensic speaker profiling. In Patrick, Schmid & Zwaan (Eds.), *Language Analysis for the Determination of Origin: Current Perspectives and New Directions*: 91–116. Cham: Springer.

Foulkes, Paul & Wilson, Kim (2011). Language analysis for the determination of origin: An empirical study. *Proceedings of the 17th International Congress of Phonetic Sciences*, 691–694. Hong Kong.

Fraser, Helen (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*(1), 113–138. DOI: 10.1558/ijsll.v16i1.113.

Fraser, Helen (2011). The role of linguists and native speakers in language analysis for the determination of speaker origin: A response to Tina Cambier-Langeveld. *International Journal of Speech, Language and the Law*, *18*(1), 121–130. DOI: 10.1558/ijsll.v18i1.121.

Goodman, Jonathan, Crema, Enrico, Nolan, Francis, Cohen, Emma & Foley, Robert (2021). Accents as honest signals of in-group membership. *Cambridge Open Engage*. DOI: 10.33774/coe-2021-c9zhv.

Haidt, Jonathan (2012). *The Righteous Mind: why good people are divided by politics and religion*. New York: Pantheon Books.

Hedegard, Hannah (2015). *Language Analysis for the Determination of Origin (LADO): The Native Speaker Dimension*. MSc dissertation, University of York.

Hoskin, James Alexander (2018). Native speaker non-linguists in LADO: an insider perspective. In Nick (Ed.), *Forensic Linguistics: Asylum-seekers, Refugees and Immigrants* (pp. 23–40). Malaga: Vernon Press.

Hoskin, James Alexander, Cambier-Langeveld, Tina & Foulkes, Paul (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. DOI: 10.1558/ijsll.39208.

Hubbuch, Priska (2019). "We Only Want to Talk…": Lingua Interviews for Linguistic Analyses for the Determination of Origin (LADO). In Patrick, Schmid & Zwaan (Eds.), *Language Analysis for the Determination of Origin: Current Perspectives and New Directions* (pp. 41–59). Cham: Springer. DOI: 10.1007/978-3-319-79003-9_5.

Hughes, Vince, Trudgill, Peter & Watt, Dominic (2012). *An Introduction to Social and Regional Varieties of English in the British Isles*. 5th Edition. London: Routledge.

International Association for Forensic Phonetics and Acoustics (2009). *Language and Determination of National Identity Cases Resolution*. Available at iafpa.net/about/resolutions/ (accessed 30 July 2023).

International Dialects of English Archive (n.d.). *England*. Available at dialectsarchive.com/england (accessed 25 May 2024).

Kinzler, Katherine D. & De Jesus, Jasmine M. (2013). Children's sociolinguistic evaluations of nice foreigners and mean Americans. *Developmental Psychology*, *49*(4), 655–664. DOI: 10.1037/a0028740.

Kinzler, Katherine D., Dupoux, Emmanuel & Spelke, Elizabeth S. (2007). The native language of social cognition. *Proceedings of the National Academy of Sciences of the United States of America*, *104*(30), 12577–12580. DOI: 10.1073/pnas.0705345104.

Liberman, Zoe, Woodward, Amanda L. & Kinzler, Katherine D. (2017). The Origins of Social Categorization. *Trends in Cognitive Sciences*, *21*(7), 556–568. DOI: 10.1016/j.tics.2017.04.004.

Matras, Yaron (2018). Duly verified? Language analysis in UK asylum applications of Syrian refugees. *International Journal of Speech, Language and the Law*, *25*(1), 53–78. DOI: 10.1558/ijsll.35710.

Matras, Yaron (2021). Process, tools and agendas in LADO: a rejoinder. *International Journal of Speech, Language and the Law*, *28*(2), 233–250. DOI: 10.1558/ijsll.20224.

Muhammad, Umar Gombe (2021). *A Comparative Phonological Analysis of Varieties of English Spoken by Native Speakers of Nigerian Languages (Hausa, Igbo, Kanuri and Yoruba) for the Determination of Speakers' Origins*. PhD thesis, University of York.

Nettle, Daniel & Dunbar, Robin Ian MacDonald (1997). Social Markers and the Evolution of Reciprocal Exchange. *Current Anthropology*, *38*(1), 93–99.

Nolan, Francis (2012). Degrees of freedom in speech production: An argument for native speakers in LADO. *International Journal of Speech, Language &the Law*, *19*(2), 263–289. DOI: 10.1558/ijsll.v19i2.263.

Patrick, Peter (2010). Language Variation and LADO. In Zwaan, Muysken & Verrips (Eds.), *Language and Origin: The Role of Language in European Asylum Procedures: Linguistic and Legal Perspectives* (pp. 73–87). Nijmegen: Wolf Legal Publishers.

Patrick, Peter (2012). Language analysis for determination of origin: Objective evidence for refugee status determination. In Tiersma & Solan (Eds.), *The Oxford Handbook of Language and Law* (pp. 533–546). Oxford: Oxford University Press. DOI: 10.1093/oxfordhb/9780199572120.013.0039.

Pinker, Steven (2012). *The False Allure of Group Selection*. Available at edge.org/conversation/steven_pinker-the-false-allure-of-group-selection (accessed 10 July 2023).

Polanyi, Michael (1966). *The Tacit Dimension*. London: Routledge & Kegan Paul.

R Core Team (2022). *R: A Language and Environment for Statistical Computing*. R Foundation for Statistical Computing. Vienna, Austria. Available at r-project.org/.

Richerson, Peter J. & Boyd, Robert (2010). Why Possibly Language Evolved. *Biolinguistics* 4(2–3), 289–306. DOI: 10.5964/bioling.8793.

Schumacher, Ernst F. (1977). *A Guide for the Perplexed*. Abacus: London.

Shen, Chen & Watt, Dominic (2015). Accent categorisation by lay listeners: which type of "native ear" works better? *York Papers in Linguistics* 2(14), 106–129.

Shorrocks, Graham (1998) *A Grammar of the Dialect of the Bolton Area*. Frankfurt am Main: Peter Lang.

Verified (2012). *LOID – Method and Limitations*. Available at cms.qz.com/wp-content/uploads/2015/09/metod-och-begrc3a4nsningar-loid-ho.pdf (accessed 13 May 2024).

Wagner, Laura, Clopper, Cynthia B. & Pate, John K. (2014). Children's perception of dialect variation. *Journal of Child Language*, 41(5), 1062–1084. DOI: 10.1017/S0305000913000330.

Watson, Kevin & Clark, Lynn (2013). How salient is the NURSE/SQUARE merger? *English Language and Linguistics*, 17(2), 297–323. DOI: 10.1017/S136067431300004X.

Weinreich, Uriel, Labov, William & Herzog, Marvin I. (1968). Empirical Foundations for a Theory of Language Change. In Lehmann & Malkiel (Eds.), *Perspectives on Historical Linguistics* (pp. 95–195). Austin: University of Texas Press.

Wells, John Christopher (1982). *Accents of English, Vol. 2: The British Isles*. Cambridge: University Press.