

The sociolinguistics of variety identification and categorisation: free classification of varieties of spoken English amongst non-linguist listeners

Robert M. McKenzie*

English Language and Linguistics, Department of Humanities, Northumbria University, Newcastle Upon Tyne, United Kingdom

*Email: robert.mckenzie@northumbria.ac.uk

In addition to the examination of non-linguists' evaluations of different speech varieties, in recent years, sociolinguists and sociophoneticians have afforded greater attention towards the ways in which naïve listeners' perceive, process and encode spoken language variation, including the identification of language varieties as regionally or socially localised forms. The present study attempts to extend understanding of non-linguists' perceptions of linguistic diversity through the investigation of how accurately and consistently UK-born students, resident in the north-east of England, can identify the speaker place of origin of six forms of L1 and L2 English. The results demonstrate that whilst the process of encoding indexical properties to and categorisations of speech stimulus as belonging to a specific language variety is complex, there is a clear tendency amongst informants to initially identify the speech as either native or non-native, most especially through the perception of specific segmental and non-segmental phonological features, before attempting more fine-grained classifications. The findings also point to the recognition of speaker place of origin at different levels of awareness, above and below the level of individual consciousness.

Keywords: speech perception, variety identification, explicit vs. implicit linguistic processing, sociophonetics, social cognition, variationist sociolinguistics

Introduction

Historically, the majority of researchers working in the area of variationist sociolinguistics have tended to focus almost exclusively upon individuals' language use. Thus, besides research investigating social evaluations of speech, conducted principally by language attitude and perceptual dialectology researchers, sociolinguistic studies have most frequently focussed upon the ways in which speakers employ language differently according to differences in their gender, age, socio-economic status, group affiliation or place of origin. Until relatively recently, the study of speech perception, i.e., how listeners' process and assign social information to the speech they hear (rather than how speech is evaluated), has been less prominent within sociolinguistics (see Campbell-Kibler, 2010). The result, as Clopper & Pisoni (2007a, p. 315) note is that despite '...large amounts of evidence to support the notion that linguistic variation between talkers due to regional and ethnic differences is real and robust and an important property of spoken language....we know less about what naïve listeners know about these sources of variation'.

However, in recent years, sociolinguists and sociophoneticians have started to more fully appreciate the value of investigating the perception of linguistic variation as a key dimension in the building of sociolinguistic theory. The findings from speech perception studies can, for instance, provide valuable information regarding the robustness of non-linguist listeners' perceptual labels of specific social and regional varieties as well as help identify the (socio)linguistic features which determine their classification and, in turn, shed greater light upon the extent to which researchers' own categorisations of language varieties map on to those of the speech communities under consideration (for an overview see Thomas, 2002). Researchers currently investigate a broad range of sociolinguistic and sociopsychological issues related to the perception of spoken linguistic diversity, most frequently through the presentation of speech stimulus comprising a series of words and/or short phrases to listeners, including: the extent to which perceived personal traits of individual speakers are based upon voice; how speakers of different languages or language varieties classify and label the same sounds; listeners' ability to perceive vowel mergers and splits; the speech perception abilities of special listeners; the influence of hearers' stereotypes on the perception of sounds; and the ability of listeners' to identify and subsequently categorise the language variety spoken and/or the geographical origin of the speaker.

In terms of the latter, there exists a series of studies investigating whether, and if so, to what extent, listeners can identify the geographical origin of speaker(s) and/or the variety spoken, based solely upon speech stimulus. The vast majority have involved measuring listeners' identifications of *L1 varieties* of a given language.

One of the earliest studies, conducted amongst native English speakers in the United States (Preston, 1993), involved the presentation of spontaneous speech recordings of 9 males to test the ability of non-linguists to discriminate between different varieties of US English. It was discovered that whilst listeners were generally able to make distinctions between Northern and Southern forms of US English - they had greater difficulty distinguishing between Northern and Midland US English - which suggests, with regards to linguistic diversity in US English, the north-south distinction is especially salient for the general public in the United States (see also McKenzie & Osthus, 2011). Further research focussing specifically upon listeners' ability to identify and categorise the place of origin of *native* speakers of the language under consideration has since been undertaken in a range of countries including: Wales (Williams, Garrett & Coupland, 1999); The Netherlands (Van Bezooijen & Gooskens, 1999); France (Boughton, 2006); Denmark (Ladegaard, 2001); Japan (Morris, 2010); and continues apace in the United

States (Purnell, Idsardi & Baugh, 1999; Thomas, Lass & Carpenter, 2010). Taken together, the results of these studies reveal that native speaker participants, asked to listen to speech stimulus, are generally able to accurately and consistently identify speakers' places of origin and/or varieties of a given language as regionally or socially localised forms, provided the regional identity of each individual speaker is not too fine-grained for the hearer. Moreover, some evidence exists that listener region of origin and level of geographical mobility can influence accuracy rates, with a tendency for more correct categorisations as well as a higher level of perceptual distinctiveness amongst listeners asked to identify varieties which are 'localised' for them and also amongst those participants who travel most extensively (Clopper & Pisoni, 2006).

Some researchers have extended the investigation to native speaker categorisations of L2 speech forms. Examples of studies measuring listener categorisations of L2 speech and which presented actual speech stimulus include: native French speaker perceptions of the L2-accented French of L1 speakers of Arabic and Western European languages (Vieura, de Mareuil & Adda-Decker, 2011); and identifications of the ethnicity of Korean speakers of English amongst L1 English speakers in the United States (Lindemann, 2003). The findings from the above studies show that whilst listeners' tend to find the categorisation task difficult, identification rates of L2 speech were generally above chance.

It is surprising that speech perception studies involving the presentation of *both* L1 and L2 varieties are rare. One early study, conducted by Stephan (1997), investigated German students' perceptions of a range of L1 English forms, as well as speakers of Nigerian English and Indian English who learned the language as an L2. The results demonstrated listeners were most able to correctly identify US and UK varieties of English, followed by the L2 forms, with Australian, New Zealand and South African English speech the least accurately identified. Stephan concluded that although the relatively high identification rates for the (standard) UK and US varieties resulted from classroom learning, differences in recognition rates between the other varieties of English speech presented were not always related to levels of previous exposure to those forms in an educational context. A recent study, conducted in a UK context, investigated the perceptions of Scottish, English and immigrant Polish adolescents of urban varieties of UK English as well as Polish English (Clark & Schlee, 2010). It was found that although recognition of the geographical origin of the Polish speakers was generally high for all listeners, the Polish participants demonstrated considerable difficulty classifying the city of origin of the urban UK varieties, most especially when the variety was spoken outwith the area of the UK in which they had settled.

In a large-scale study, McKenzie (2008a) investigated how accurately Japanese-born university students' could identify the place of origin of four speakers of standard and non-standard varieties of Scottish and US English and two speakers of Japanese English. The results demonstrated that the provenance of a heavily-accented speaker of Japanese English was the most accurately identified (over 90%), followed by the speakers of US English (55% and 59%), with the speakers of Scottish English (32% and 31%) and a moderately-accented speaker of Japanese English (30%) the most difficult for listeners' to categorise. Follow-up comments also indicated that aspects of a speaker's pronunciation, most especially segmental features (e.g., the rhoticity of the Scottish English speech), rather than morpho-syntactic or lexical features, were responsible for the listeners' identifications and misidentifications. McKenzie also found evidence that the Japanese students' were better able to correctly categorise the speakers as either native or non-native and concluded the distinctions made between L1 and L2 speech constituted a salient initial stage for the listeners to perceive, process, encode and ultimately classify the variation in the speech stimulus presented.

It is notable that previous studies measuring hearers' ability to discriminate between language varieties have tended to employ forced-choice categorisation tasks, i.e., involving the presentation of a closed-set of potential countries and/or language varieties for participants to choose from when listening to the speech stimulus. However, the use of an unconstrained free classification task, where listeners' are able to provide their own labels, has a number of advantages over a forced-choice task. First, the constraints imposed by the researchers in a forced-choice task, by its very design, i.e., providing a limited selection of countries or speech varieties, restricts listeners from creating their own categorisations (Clopper & Pisoni, 2007b). For this reason, allowing free classifications is likely to offer more information regarding the depth of listeners' linguistic awareness of the speech varieties under consideration where, for instance, the classification of a speaker of Scottish Standard English as either from the 'UK', 'Scotland' or 'Glasgow' would seem to indicate the extent to which the listener can pinpoint where the variety is spoken, and thus, help reveal the depth of his/her level perceptual categorisations of the speech. Moreover, since variety classifications tend to be based upon the ethnic associations of the listener, and perhaps especially when the speech stimulus includes L2 forms which native speakers of the language in question have had less exposure to (McKenzie, 2008a), analysis of the participants' own labels may provide more detailed information. For example, in the case of listeners who incorrectly identify the stimulus speech as a linguistic variety with which they are more familiar (and which they may associate/conflate with the

misidentified variety), analysis of the patterns of identification and misidentification may provide an insight into the stereotypes and wider ideological frameworks of the participants (McKenzie, 2010), for instance, by providing evidence of the extent to which L1 speakers of English in the UK conflate different forms of English spoken in North America or East Asia.

The present study attempts to extend our understanding of non-linguists' perceptions of linguistic diversity through the investigation of how accurately and consistently UK-born, native English speaking university students can identify speaker place of origin of specific varieties of L1 English varieties spoken in the UK, as well as a range of L2 English forms spoken by overseas students at the university in question (see below). It is envisaged that the findings gained from the study may also provide information into how listeners use their knowledge of linguistic variation in order to identify the regional provenance of a given speaker. It is also hoped that analysis of the patterns of identification and misidentification may offer some indication of the precise linguistic features (or more likely, interactions between features) upon which listeners base their classifications upon as well as reveal more about the participants wider ideological frameworks concerning linguistic diversity in English.

Method

Participants

The sample was composed of 194 UK-born nationals studying at Northumbria University, situated in the city of Newcastle in the north-east of England. It was decided to recruit university students specifically as participants since it was considered that the young, mobile, educated individuals, who tend to constitute this group, were most likely to be exposed to the widest range of diversity in English language and, given the recent unprecedented rise in numbers of overseas students attending British higher educational institutions, much more likely to be familiar with Asian forms of English than other sections of UK population (see below). Since all of the students were undergraduates, the overwhelming majority of the listeners were aged between 18 and 21 (mean = 20.2, *sd* = 4.6). The responses of a small number of listeners who did not report their place of birth as the UK and/or as native speakers of English were discarded. Hence, the sample appeared to be composed solely of university students of UK nationality, who spoke English as an L1 and, at the time of the data collection, lived in the north-east of England. Indeed, the great majority were born and raised in the region.

Speech Stimulus

The stimulus consisted of 6 samples of spontaneous English. From a larger database of high quality digital-audio recordings, individually recorded by the researcher, each of the 6 samples was validated as most representative of that particular form of English by means of a pilot study involving the judgments of different focus groups of fluent speakers of the variety (it is acknowledged that the UK-born participants involved in the present study may potentially have selected different recordings as representative of each variety). To minimise speaker gender as a potentially confounding factor, the recordings were provided by six female speakers of English. To further ensure control over other extraneous variables each of the speakers was recorded giving directions on a fictitious map, by means of a map-task (see McKenzie, 2008b), thus guaranteeing that no references were made regarding the geographical origin of the speaker or the variety of English spoken. Speakers of the following varieties were employed for the purposes of identification (relevant speaker details are also included):

Scottish Standard English (ScotStE): 86 seconds. L1 English. Age 30. Born near Glasgow, Scotland. Describes her English as ‘soft Scottish English’.

Tyneside (Newcastle) English (TyneE): 67 seconds. L1 English. Age 23. Born Newcastle, England. Describes her English as ‘Geordie’.

Indian English (IndE): 64 seconds. L1 Tamil. L2 English. Age 27. Born Tamil Nadu, South India. Describes her English as ‘India accent’.

Japanese English (JapanE): 89 seconds. L1 Japanese. L2 English. Age 33. Born Hyogo, Japan. Describes her English as ‘English with Japanese accent’.

Chinese English (ChinE): 82 seconds. L1 Putonghua Chinese. L2 English. Age 24. Born Xian, China. Describes her English as ‘Chinese style but speak standard’.

Thai English (ThaiE): 88 seconds. L1 Thai. L2 English. Age 25. Born Chiang Rai, Thailand. Describes her English as ‘English with Thai features’.

The choice of these varieties as speech stimulus was quite deliberate. It was felt that for

listeners from Newcastle, Tyneside English and Scottish Standard English together represent the local variety of English and, arguably, for individuals living in a town approximately 35 miles from the Scottish border, the most socially salient regional standard variety of English, i.e., the variety which carries the greatest weight of social indexation as a regional standard (see Racz, 2013). The four Asian forms of English speech were selected precisely because, at the time of the data collection, overseas students from India, Japan, China and Thailand constituted the largest groups of international students attending the university where the data was collected and hence, it was considered most appropriate to investigate home students' familiarity with these spoken forms of English. Since all four Asian speakers had learned English as an L2, to ensure that listeners were responding to linguistic differences between the speech varieties and not English language proficiency, each of these individuals had attained an advanced level in the language and, indeed, had obtained a university degree taught in English prior to the recordings.

Research Instrument

In accordance with previous variety recognition research (see McKenzie, 2008a; 2010), to ascertain the extent to which the listeners could correctly and consistently identify the place of origin of each of the speakers, the participants listened to each of the 6 speech samples and were then requested to provide written responses to the question 'which country do you think the speaker comes from?' A follow-up question, 'how did you make this decision?' was also incorporated into the design of the research instrument, with the aim of gaining as much fine-grained information as possible regarding the reasons for the participants' classifications of the speakers' place of origin.

Procedure

The variety identification task was carried out amongst groups of students at the university in question in late 2011 and early 2012. Participants were allowed to hear each of the speech varieties once only.

Results and Discussion

Initial analysis involved the calculation of recognition rates for the provenance of each of the 6 speakers (Table 1). Given the often idiosyncratic nature of the responses provided these calculations were not always straightforward. Although listeners were requested to name a specific country of origin for each speaker, a number of listeners, in particular, employed

terminology such as ‘East Asia’, ‘Asia’ and, most especially, ‘non-native speech/speaker’ to classify the Thai, Japanese and/or Chinese speakers of English (see also below). In the case of the latter, whilst the labelling of an individual as either a native speaker or a non-native speaker remains controversial amongst some academics (see Davies, 2013; McKenzie, 2013), given the listener responses, a decision was taken to also calculate recognition rates for the categorisation of the six speech speakers as either ‘native’ or ‘non-native’ (Table 2).

Table 1 and Table 2 below demonstrate that there were enormous differences between the recognition rates for the perceived place of origin of the six speakers. Analysis and detailed discussion of the students’ identification and misidentification rates, together with example follow-up comments for each of the English speech forms presented, are detailed below. To provide greater clarity, separate tables detailing patterns of identification and misidentification for each of the varieties are also provided (Tables 3-8).

Table 1. Percentages (and frequencies) of correct and incorrect identifications for speaker country of origin (N=194)

Recognition	Speaker					
	ScoStE	TyneE	IndE	JapanE	ChinaE	ThaiE
Correct	96.9 (188)	95.9 (186)	93.3 (181)	26.3 (51)	11.9 (23)	6.7 (13)
Incorrect	3.1 (6)	4.1 (8)	6.7 (13)	73.7 (143)	88.1 (171)	93.3 (181)
Total	100 (194)	100 (194)	100 (194)	100 (194)	100 (194)	100 (194)

Table 2. Percentages (and frequencies) of correct and incorrect categorisations for native/non-native speaker (N=194)

Recognition	Speaker					
	ScoStE	TyneE	IndE	JapanE	ThaiE	ChinaE
Correct	99.5 (193)	99.5 (193)	99.5 (193)	99.5 (193)	94.8 (184)	87.6 (170)
Incorrect	0.5 (1)	0.5 (1)	0.5 (1)	0.5 (1)	5.1 (10)	12.4 (24)
Total	100 (194)	100 (194)	100 (194)	100 (194)	100 (194)	100 (194)

Tyneside English and Scottish Standard English

The Tyneside English (95.9%) and Scottish Standard English (96.9%) speech were clearly the most accurately and consistently identified. It is reasonable to assume that the British-born students, perhaps unsurprisingly, were most familiar with regional and social variation in English within (the north of) the UK. Participants who were able to identify the place of origin for the speaker from the north-east of England generally classified her to be from ‘Newcastle’ or ‘Tyneside’ or as a ‘Geordie’. Many commented upon their general familiarity with and/or the distinctiveness of the speech. This is perhaps unsurprising considering, as stated above, all the participants lived in the area during the data collection period, and many were also from the Newcastle area:

‘typical Geordie. Friendly, confident accent. Distinctive tone’ (71)

‘rough, thick accent and stands out anywhere’ (104)

‘clear English speaking, recognisable as a regional accent’ (112)

Those listeners who provided comments regarding the specific linguistic features which aided their correct identification of the speaker’s provenance overwhelmingly focussed upon

specific phonological features of the sample as indexical of Tyneside speech, most especially

the pronunciation of ‘l’, ‘t’, ‘c’, ‘p’ and ‘r’:

‘accent – pronunciation of “t”, “l” and vowels. Very fluent native English speaker’ (65)

‘omission of “t”’ (63)

‘stereotypical accent not using letter t and c and p (83)

‘accent sounds like me! eg forward – misses r’ (9)

These comments seem to reflect the recognition of specific features of spoken Tyneside English, all present in the speech sample employed: clear (i.e., non-velarised) /l/ in all positions, the lack of rhoticity in comparison with other varieties spoken in the north of England, the localised glottal realisation of /p, t, k/ as /ʔp, ʔt, ʔk/ as well as glottal replacement of /t/ (in the case of the latter, a feature increasingly employed by younger speakers in many urban centres throughout the UK) (see Milroy, Milroy, Hartley & Walshaw, 1995; Kerswill, 2003).

Table 3. Tyneside English: Percentages (and frequencies) of perceived speaker origin (N=194)

Speaker Origin	Percentage (Frequency)
Newcastle/Tyneside	95.9 (186)
Other England	3.6 (7)
Unsure	0.5 (1)
Total	100 (194)

It is worth mentioning that the sample of Tyneside speech employed in the present study was provided by a younger female who, although born and raised in the city, was not found to employ any specific lexical or morpho-syntactic features associated with the variety. The small number of participants who misidentified the Tyneside speaker (3.6%) all recognised the provenance of the speaker as England, suggesting some degree of recognition. For example:

‘Similar accent to Geordie one, not quite as accented, however’ (English Midlands) (160)

Likewise, the high percentage of listeners who accurately identified the Scottish Standard

English speaker's place of origin, i.e., as Scotland/Glasgow, frequently highlighted particular phonological features, most especially the rhoticity of the speech. Examples of this rhoticity present in the speech stimulus, and mentioned by the listeners, include 'church', 'turn' and 'sharply'. Several participants also mentioned the distinctive vowel sounds, most likely reflecting the lack of systematic durational differences between long and short vowels in SSE and other varieties of Scottish English which, in contrast, are present in most other forms of English spoken in the UK:

'the sound, rolling r's. Also very fluent so learnt English as a first language' (53)

'sound of the r and vowels' (138)

'intonation on the vowels' (68)

Many participants also noted the use of Scottish lexis, most especially 'wee' and/or 'wee kink' as typically Scottish:

“‘wee kink’ narrows down to Scottish/Irish but she doesn't sound Irish' (103)

'I have been taught by many Scottish teachers so the accent is familiar. Plus the use of “wee” is a word I associate with Scotland' (133)

Some participants' categorised the provenance of the SSE speaker, born and raised in Glasgow in the west of Scotland, as 'Edinburgh'. This pattern of misidentification is intriguing since it implies for many northern English students either, that Edinburgh, as the Scottish city closest to Newcastle, is stereotypical of Scotland and in turn, of Scottish English generally, or more likely, that as the capital city and most prosperous area of Scotland, Edinburgh is associated specifically with *prestige* features of Scottish Standard English:

'intonation at end of phrases –accent not as strong as expected from Scottish' (Edinburgh, 57)

Again, the relatively low proportion of UK students who failed to accurately identify the country of origin of the Scottish English speaker (3.1%) overwhelmingly categorised her correctly as an L1 user of English:

'the voice rises sharply at points creating the accent' (Ireland) (38)

‘strong accent gives it away’ (northern Geordie accent) (161)

Table 4. Scottish Standard English: Percentages (and frequencies) of perceived speaker origin (N=194)

Speaker Origin	Percentage (Frequency)
Scotland	96.9 (188)
Other UK	2.6 (5)
L2 Europe	0.5 (1)
Total	100 (194)

Indian English

A relatively high percentage of participants were also able to identify the provenance of the speaker from India (93.3%), and overwhelmingly categorised Indian English as non-native speech (99.5%). This finding indicates that the distinction between native and non-native speech is salient for UK-born listeners when the speech of both L1 and L2 users of English is employed as stimulus. A number of listeners remarked that they had prior exposure to similar speech, for instance, through interactions with immigrants of South Asian heritage in Indian restaurants in the UK, or through conversations with call-centre operators in India:

‘People in Newcastle Indian restaurants speak like this’ (95)

‘sounds like HSBC call centre’ (121)

Specific linguistic features were also identified, again relating mainly to phonological features, particularly the realisation of ‘th’, i.e., /θ/ and /ð/, which does not exist in Indian English, and most often replaced by the (non-Dravidian) Indian dental plosives /t̪/ /d̪/, and /t̪ʰ/ /d̪ʰ/:

‘from the TH sounds – the way is pronounced by the speaker is a characteristic feature’ (90)

Similarly, a number of participants also noted the distinctive pronunciation of ‘d’ and ‘t’: the most plausible explanation suggests an awareness of retroflex /t̪/ and /d̪/ in the articulation of first sounds in lexis such as in ‘turn’ and ‘demand’ by many Indian speakers of English (see Sailaja, 2009), including in the Indian English speech sample included in the present study:

‘common or typical of Indian accent sound of t and d’ (55)

‘rolling of R and very pronounced T and D’ (65)

More negatively, many listeners felt they were able to correctly classify the provenance of speaker as Indian precisely because the speech was evaluated as incorrect English, especially regarding the speaker’s use of specific grammatical, phonological and/or paralinguistic features:

‘missed out words e.g “the”, incorrect English. Familiar accent I have heard before’ (126)

‘turning left –not turn left’ (124)

‘pronunciation of words and tone is not entirely fluent’ (61)

‘long pause in speech, stuttering’ (57)

The relatively low proportion of listeners who failed to identify the provenance of the Indian speaker often classified her as either Bangladeshi, Pakistani or Sri Lankan, providing compelling evidence that whilst northern English university students possess stereotypes of English spoken in the wider Indian sub-continent, some may face more difficulties distinguishing between forms of the language spoken in the countries/regions within this large area:

‘Similar to Pakistan accents I have heard’ (Pakistan) (168)

Table 5. Indian English: Percentages (and frequencies) of perceived speaker origin (N=194)

Speaker Origin	Percentage (Frequency)
India	92.8 (180)
Other Indian Subcontinent	2.6 (5)
Other L2 (Africa, Central America, not specified)	1.5 (3)
L2 Europe	1.5 (3)
Unsure	0.5 (1)
Other Asia	0.5 (1)

L1 (UK)	0.5 (1)
Total	100 (194)

Japanese English

The task of identifying the provenance of the 3 East Asian speakers was clearly more problematic. For example, although the Japanese speaker of English was identified most frequently of the three East Asian speakers included in the study, only 26.3% classified her country of origin correctly. Listeners who correctly recognised the provenance of the Japanese speaker tended to comment upon the speech as non-native, and, in turn, highlighted a number of ‘incorrect features’, such as the perceived errors in pragmatic use the speaker made, the lack of opposition between /l/ and /r/ and a general lack of fluency:

‘Use of word the English wouldn’t necessarily use in this context eg repetition of please’ (93)

‘the use of L in place of R and the use of word please after sentences. Not native speaker’ (18)

‘not entirely fluent, difficulty to pronounce certain words’ (61)

Several participants also noted the high pitch of the speaker’s voice, a feature of the formal speech of female speakers of Japanese:

‘have Japanese friend who has high pitched tone like this’ (95)

A number of listeners commented more positively upon the politeness of the speaker, most especially through the use of ‘please’ whilst giving directions:

‘use of word please may times. Very polite culture. Accent seems familiar’ (99)

‘Politeness, pronunciation of r’ (67)

The larger percentage of participants who were unable to identify the country of origin of the speaker (73.7%), nonetheless, were frequently able to identify the speech as non-native English (99.5%), providing further evidence that the distinction between native and non-native English is marked for the listeners. Many of these participants classified the speech as L2 English precisely because it was perceived as incorrect:

‘Words left out of the sentence. Suggests English is not their first language’ (Korea) (60)

It is notable that the vast majority of listeners who were unable to accurately identify the provenance of the Japanese speaker misidentified her to be from another country in East Asia, often China. Several participants mentioned the lack of opposition between /l/ and /r/ and frequent (mis)use of polite speech as indexical of East Asian forms of English, mirroring comments made by many participants who were able to identify the speaker correctly:

‘pronunciation l for r and r for l’ (China) (122)

‘A clear Chinese accent –repeated use of word “please”, I think is common with Chinese speakers’ (China) (111)

‘Repetition of please and not in the right context’ (South Korea) (67)

Table 6. Japanese English: Percentages (and frequencies) of perceived speaker origin (N=194)

Speaker Origin	Percentage (Frequency)
Other East Asia	48.5 (94)
Japan	25.8 (50)
L2 Europe	15.5 (30)
Other Asia	6.7 (13)
Unsure	2.1 (4)
Other L2 (Africa, South America)	1.5 (3)
Total	100 (194)

Chinese English and Thai English

The UK-born students demonstrated considerable difficulty in terms of accurate identifications of the country of origin of the Chinese speaker (11.9%) and the Thai speaker (6.7%). In the case of the Chinese speaker, the relatively low proportion of students who were able to correctly identify the speaker’s provenance tended to remark upon segmental features of the speech presented. A number highlighted the lack of opposition between /d/ and /θ, ð/, a feature shared with other L2 English forms, as well as a very particular use of ‘r’ amongst Chinese speakers of English, perhaps reflective of their previous interactions with speakers of English from China as the

largest cohort of overseas students attending the university:

‘sounds similar to my friend when he arrived from China, sound of word road is telling’ (182)

‘pronunciation of th’ (115)

‘r pronounced strongly’ (129)

A few listeners commented upon some similarities between the phonetic features of the sample of Chinese English speech presented and Putonghua Chinese:

‘I am familiar with this accent. Some of the pronunciation is similar to mandarin Chinese’ (181)

A number of participants, however, did not provide any comments even when they could accurately identify the place of origin of the Chinese speaker, seemingly indicating a broad lack of *conscious* (i.e., explicit) awareness of the linguistic features of the English spoken in China.

The much higher percentage of participants who could not identify the Chinese speaker’s place of origin, nonetheless, were again generally able to recognise the speech as non-native (87.6%). A sizeable proportion (15.9%) correctly classified the speaker to be from (East) Asia but misidentified the specific country:

‘a little Korean’ (Korea) (136)

‘not fluent’ (Philippines) (45)

Intriguingly, and contrary to expectation, a substantial proportion (47%) misidentified the Chinese speaker to come from a specific country in (mainly northern or eastern) Europe or to be an L2 English speaker from an unspecified country in ‘Europe’ more generally (i.e., they were unwilling/unable to identify a European country):

‘Fairly fluent but slight accent “when you get to the airport” – missing out “to”’ (Swedish) (74)

‘not first language, from Europe somewhere’ (Europe) (190)

‘the broken speech is what I associate with Eastern European accents (Eastern Europe) (133)

The above comments suggest the Chinese speaker was initially identified as an L2 user of English before more fine-grained (mis)categorisations were made and again, for many, the pronunciation of the speaker appeared to be most perceptually distinctive feature for their choice. The comments nonetheless indicate that participants were generally unable (or unwilling) to identify any specific phonological or phonetic features responsible for their classifications:

‘not fluent, yet their pronunciation is very good’ (France) (44)

‘well spoken English with a hint of accent’ (northern Europe) (153)

One potential explanation for this somewhat surprising pattern of misidentification points to specific linguistic features (e.g., voice quality) of the English spoken and/or non-linguistic features (e.g., personality) of the individual chosen for the study to be responsible for the categorisations. However, this explanation seems unlikely since, as described above, in the present study, all the English speech stimuli employed, including the sample provided by this Chinese national, were validated as most representative of Chinese English in the earlier pilot study (see also McKenzie, forthcoming).

Table 7. Chinese English: Percentages (and frequencies) of perceived speaker origin (N=194)

Speaker Origin	Percentage (Frequency)
L2 Europe	47.4 (92)
Other L2 (Africa, South America, not specified)	12.9 (25)
China	11.9 (23)
Unsure	10.8 (21)
Other East Asia	8.2 (16)
Other Asia	7.7 (15)
L1 (Australia)	1.0 (2)
Total	100 (194)

Listeners clearly had most difficulty accurately identifying the geographical origin of the Thai speaker, with a mere 6.7% recognition rate. Those participants who correctly identified the speaker’s provenance frequently did so as a result of perceived errors in the speaker’s use of

English, most especially regarding perceived lapses in pronunciation and grammar and a general lack of fluency:

‘accent, poor English’ (117)

‘speaks slower, tenses are sometimes wrong’ (164)

Some listeners commented upon the distinctive intonation of the Thai speaker’s English. Indeed, Thai is a tonal language and, as in the sample employed in the present study, there is a tendency amongst Thai speakers of English to assign tone to individual syllables (Trakulkasemsuk, 2012):

‘there are different tones in her speech’ (151)

A substantial percentage of listeners who failed to recognise the Thai speaker’s country of origin were again able to identify the speech as non-native (93.35%), and to come from either another South-East Asian nation (7.2%) or, more often, from elsewhere in the continent (60.3%):

‘Fluent but the grammar and words are not British’ (Hong Kong) (112)

‘sounds Chinese, though somewhat more accented’ (Korea) (160)

Listeners who misidentified the provenance of the Thai speaker frequently remarked upon perceived pronunciation errors, in relation to either segmental or prosodic features, including the realisation of /ð/and/or /t/ as /d/, lack of opposition between /l/ and /r/ and elision and/or lenition of consonants in word initial and word final position:

‘they say “the” as “duh”. Sentences are broken and not confidently spoken’ (‘Oriental, not sure of country’) (124)

‘Misses off ‘t’ in left, da instead of the’ (China) (123)

‘the dropping of words on last letter (lef not left) and lack of “r” sounds’ (Japan) (133)

‘Tone. Silent “v” sounds’ (Malaysia) (97)

The most likely explanation for the relatively low recognition rate for the Thai

speaker is that, because of the comparatively low numbers of Thai nationals studying in the UK, when compared to the higher numbers of students from China, Japan or India, together with the limited attention paid to Thai culture (and, in turn, Thai English) within the British media, UK-born students have lower levels of exposure to Thai English, i.e., they do not have reliable perceptual records of the norms of Thai English speech (Williams, Garrett & Coupland, 1999). Follow up comments point to the difficulty many participants clearly faced to identify the geographical origin of the speaker and also suggest a comparatively limited amount of exposure to English speech of Thai nationals:

‘guess. accent is hard to pinpoint’ (Asia) (64)

‘not confident with the language. Monotone’ (no classification offered) (50)

Table 8. Thai English: Percentages (and frequencies) of perceived speaker origin (N=194)

Speaker Origin	Percentage (Frequency)
Other Asia	60.3 (117)
Other L2 (Africa, South America, not specified)	10.8 (21)
L2 Europe	9.8 (19)
Other SE Asia	7.2 (14)
Thailand	6.7 (13)
Unsure	4.6 (9)
L1 (USA)	0.5 (1)
Total	100 (194)

Conclusion

In addition to the examination of non-linguists’ evaluations of different speech varieties, in recent years, sociolinguists have afforded greater attention towards non-linguists’ perceptions of variation in speech, including the identification of language varieties as regionally or socially localised forms. The present study investigates the extent to which UK-born university students’ are able to correctly identify speaker country of origin, together with reasons for the choices made, for 6 forms of L1 and L2 English speech. Patterns of misidentification were also examined.

Since a free variety identification task was employed in the present study it was envisaged that accurate categorisation rates would be somewhat lower than equivalent studies involving forced-choice tasks (e.g., Williams, Garrett & Coupland, 1999; Van Bezooijen & Gooskens, 1999). However, the results demonstrate that UK-born participants, resident in the Newcastle area, were generally able to accurately and consistently identify the place of origin of 'local' speakers and a speaker of Indian English, and thus, provide evidence that they possess robust perceptual categorisations of these varieties. This finding is broadly consistent with the results of the aforementioned study by Clarke and Schlee (2010), where recognition rates for both UK-born and Polish-born adolescents living in the UK, perhaps unsurprisingly, were also highest for urban forms of UK English which were most localised to them.

Listeners clearly found categorising the provenance of the speakers from Japan, China and Thailand more problematic (and indeed, in the case of the Chinese speaker in particular, were frequently unable to detail any specific features responsible for their choices). The most plausible explanation for the relatively low hit-rates, despite the presence of large groups of students from these 3 East Asian countries attending the University, is a lack of interaction with speakers of, and thus a low level of familiarity with, these L2 forms of English. Interestingly, differences between students' identification rates for all 4 Asian speakers of English (i.e., also including the Indian English speaker) cannot be explained by differences in numbers of students from each country attending the university in question.

Fine-grained analysis of the patterns of identification and misidentification demonstrated that many of the UK-born students who incorrectly identified the geographical origin of the speakers of English from Japan, China and Thailand, perceived them to be from different countries within East Asia, i.e., their perceptual categories were not sufficiently robust to pinpoint the precise country of origin (see above). The great majority of listeners were able to categorise these 3 samples as L2 English more broadly, implying the recognition of the speech at some level of awareness. Follow-up participant comments, detailing the reasons for their speaker origin classifications, also frequently highlighted perceived differences between the English of the native speakers and the non-native speakers presented for identification.

In the case of the L2 speech samples, analysis of participant comments points, rather disappointingly, to the existence of a deficit model in relation to identifications of non-native English speech more generally, with a tendency for the UK-born students to attribute their classifications of Indian, Japanese, Chinese and Thai English speech to perceived errors the speakers made, for both correct and incorrect provenance categorisations. Comments indicated

that it was phonological errors these speakers were perceived to make which were deemed most responsible for judgements of their speech as L2 English (e.g., in the lack of opposition between specific phonemes), albeit frequently in conjunction with perceived lapses in grammar and lexis, in the distinctive usage of pragmatic features and at the suprasegmental level. Although unrelated to perceptions of correctness and incorrectness, this result is broadly compatible with the findings from studies involving native speaker identifications of different L1 speech forms, where evidence was also found that it was mainly phonological cues, together with prosodic and other linguistic features, which enabled listeners to recognise speaker provenance from speech stimuli (e.g., Van Bezooijen & Gooskens 1999; Ladegaard, 2001).

The above discussion, in conjunction with the results of prior research involving the presentation of L1 speech varieties (e.g., Williams, Garrett & Coupland, 1999), suggests requesting listeners' identify the geographical origin of speakers also invites social evaluation of the speech, and in the case of the present study, of the speakers' social/ethnic group membership. This seems the case for both the L1 and L2 forms of English speech presented, since each speaker was frequently identified correctly as a native or a non-native user of the language precisely because their speech was deemed to be correct (for an L1 speaker) or incorrect (for an L2 speaker). The more positive comments relating to the local variety of English spoken by many of the students themselves, i.e., Tyneside English, reveals a level of linguistic security (see Labov, 2006), at least when the variety is presented alongside forms of Asian English speech. More negatively, the results also imply that some (Asian) L2 speech forms may represent non-standard English for many UK-born students, arguably reflecting the continuing stigmatisation of particular groups of Asians in the UK more broadly (see Ford, 2008).

Comments regarding the perceived incorrectness of Asian English speech amongst UK university students parallels the findings of a study by McKenzie (2008a) amongst Japanese university students, where the native varieties of English speech presented for identification were generally considered to embody 'notions of correctness' but the English spoken by Japanese nationals was largely stigmatised. McKenzie attributed the findings to the construction of a pervasive 'native speaker ideology' amongst Japanese users of English more widely. A preference for English perceived as 'native' over local forms of English speech has also been demonstrated in recent studies undertaken in other Asian contexts, e.g., China (Xu, Wang & Case, 2010), South Korea (Yook & Lindemann, 2013) and Oman (Buckingham, 2014).

The results of the present study also suggest that mother-tongue speakers of English, at least within the UK university context, also tend to regard the language of native speakers of

English to be the most correct. Evidence for the construction of native speaker ideologies amongst other L1 English speaking populations, especially in the United States, also exists. More specifically, the results from these language attitude point towards US-born students' evaluations of L2 English as largely based upon ethnic dimensions, whereby the more 'Caucasian' a speaker is judged to be, the more prestigious his/her speech is likely to be rated (e.g., Lindemann, 2003; Lippi-Green, 2012).

The findings of the present study demonstrate that the process of encoding indexical properties of speech and the categorising of speech forms as belonging to a particular language or language variety are clearly complex and that 'recognition' of a speech variety and/or the place of origin of an individual speaker can occur at different levels of awareness. This is perhaps unsurprising given there is evidence to suggest speech recognition, in general, requires a complex interaction between detailed featural analysis and overall pattern recognition, at both conscious and unconscious levels (Kreiman & Sidts, 2011). The results of prior variety recognition studies, involving the presentation of either L1 or L2 speech stimulus, also demonstrates the complex nature of speaker categorisation, where there is evidence to suggest the use of unconscious as well as conscious processes in speech categorisations, for instance, involving Danish learners' identifications of UK, US and Australian English (Ladegaard, 1998) and US nationals' recognition of English speech as 'foreign' (Preston, 1996). In the present study, fine-grained analysis of patterns of misidentification and the detailing of follow-up participant comments, in addition to the calculation of recognition rates, provides particularly compelling evidence that the UK-born students involved in the study are also able to employ more than linguistic knowledge which is 'overtly available' (ibid) in order to discriminate between different forms of L1 and L2 speech and, in turn, to identify the provenance of speakers.

It is felt that the findings of the present study shed additional insight into the complex nature of non-linguists' ability to differentiate between and to classify different varieties of L1 alongside L2 English speech. However, equivalent studies involving the presentation of both native and non-native speech stimuli are relatively limited, most especially involving forms of Asian English speech stimulus and amongst UK nationals as listeners. Undoubtedly, there is much scope for future research. For instance, participants in the present study were composed exclusively of UK-born university students attending one institution in the north-east of England. Clearly, in order to generalise the findings beyond this particular group, it would be profitable to replicate the study with other populations, such as adolescents and older adults in different areas as well as amongst different ethnicities, both within and outwith the UK. Likewise, in future

studies it would be of value to investigate whether any differences within the population selected, e.g., in terms of gender, educational background or social mobility, can influence categorisations. There is also a clear need to present participants with recordings of speakers of other forms of L1 and L2 English for identification, spoken outwith the UK, the US, the Indian subcontinent or East Asia. Similarly, whilst the present study employed more lengthy spontaneous samples of L1 and L2 English speech as stimuli, it is only through the additional presentation of shorter fragments of manipulated speech which would help researchers understand more fully which components of the acoustic speech signal are most salient for listeners, e.g., vowels, consonants, prosodic features and/or voice quality, and thus allow for the identification of the provenance of speakers of different language varieties (see Purnell, Idsardi & Baugh, 1999 for an interesting identification study of L1 speech varieties based upon short speech fragments).

Finally, in order to determine response latency, i.e., how rapidly listeners can identify a variety, researchers have the option to adapt measurement tools to oblige participants to offer responses under time pressure (Clopper & Pisoni, 2007a). The findings obtained from such studies would help build up a more detailed picture of the underlying mental processes by which non-linguist listeners make their choices and, in particular, may help indicate the extent to which variety categorisations and identifications of speaker place of origin involve implicit and/or explicit language processing abilities, i.e., to help determine to the extent to which the ability to perceive, encode and classify variation in spoken language is above or below the level of individual consciousness.

References

- Boughton, Z. (2006). When perception isn't reality: accent identification and perceptual dialectology in French. *French Language Studies*, 16(3), 277-304.
- Buckingham, L. (2014). Attitudes to English teachers' accents in the Gulf. *International Journal of Applied Linguistics*, 24(1), 50-73.
- Campbell-Kibler, K. (2010). Sociolinguistics and perception. *Language and Linguistics Compass* 4(6), 377-389.
- Clark, L. & Schlee, E. (2010). The acquisition of sociolinguistic evaluations among Polish-born adolescents. *Language Awareness*, 19(4), 299-322.
- Clopper, C. & Pisoni, D. B. (2006). Effects of region of origin and geographical ability on perceptual dialect categorization. *Language Variation and Change*, 18(2), 193-221.
- Clopper, C. & Pisoni, D. B. (2007a). Perception of dialect variation. In D. B. Pisoni & Remez, R. E. (Eds.), *The handbook of speech perception* (pp. 313-337). Oxford: Blackwell.
- Clopper, C. & Pisoni, D. B. (2007b). Free classification of regional dialects of American English. *Journal of Phonetics*, 35(3), 421-438.
- Davies, A. (2013). *Native speakers and native users: loss and gain*. Cambridge: Cambridge University Press.
- Ford, R. (2008). Is racial prejudice declining in Britain? *The British Journal of Sociology* 59(4), 609-636.
- Kerswill, P. (2003). Dialect levelling and geographical diffusion in British English. In D. Britain

- & J. Cheshire (Eds.), *Social dialectology: In honour of Peter Trudgill* (pp. 223-243). Amsterdam: John Benjamins.
- Kreiman, J. & Sidts, D. (2011). *Foundations of voice studies*. Oxford: Wiley-Blackwell.
- Labov, W. (2006). *The social stratification of English in New York City*. (2nd ed.). Cambridge: Cambridge University Press.
- Ladegaard, H. J. (1998). National stereotypes and language attitudes: The perception of British, American and Australian language and culture in Denmark. *Language and Communication*, 18, 251-274.
- Ladegaard, H. J. (2001). Popular perceptions of standard regional language: attitudes to 'regional standards' in Denmark. *Language Awareness*, 10(1), 25-40.
- Lindemann, S. (2003). Who speaks "broken" English? US undergraduate perceptions of non-native English. *International Journal of Applied Linguistics*, 15(2), 187-212.
- Lippi-Green, R. (2012). *English with an accent*. (2nd ed.). London: Routledge.
- McKenzie, R. M. (2008a). The role of variety recognition in Japanese university students' attitudes towards English speech varieties. *Journal of Multilingual and Multicultural Development*, 29(2), 139-153.
- McKenzie, R. M. (2008b). Social factors and non-native attitudes towards varieties of spoken English: a Japanese case study. *International Journal of Applied Linguistics*, 18(1), 63-68.
- McKenzie, R. M. (2010). *The social psychology of English as a global language*. Dordrecht, Netherlands: Springer.
- McKenzie, R. M. (2013). Changing perceptions? A variationist sociolinguistic perspective on

- native speaker ideologies and standard English in Japan. In S. Houghton & D. Rivers (Eds.), *Native speakerism in Japan* (pp. 215-226). Bristol: Multilingual Matters.
- McKenzie, R. M. (forthcoming). UK University students' folk perceptions of global variation in English: the role of explicit and implicit attitudes. *International Journal of the Sociology of Language* 236.
- McKenzie, R. M. & Osthus, D.S. (2011). That which we call a rose by any other name would sound as sweet: folk perceptions, status and language variation. *AILA Review*, 24, 100-115.
- Milroy, J., Milroy, L., Hartley, S. & Walshaw, D. (1995) Glottal stops and Tyneside glottalization: competing patterns of variation and change in British English. *Language Variation and Change*, 69(3), 327-357.
- Morris, M. Y. (2010). Regional stereotypes and the perception of Japanese vowel devoicing. In D. R. Preston & N. Niedzielski (Eds.), *A reader in sociophonetics* (pp. 191-202). New York: Walter de Gruyter.
- Preston, D. R. (1993). Folk dialectology. In D. R. Preston (Ed.), *American dialect research* (pp. 333-378). Amsterdam: John Benjamins.
- Preston, D. R. (1996). Whaddayaknow?: The modes of folk linguistic awareness. *Language Awareness*, 5(1), 40-74.
- Purnell, T., Idsardi, W. & Baugh, J. (1999). Perceptual and phonetic experiments on American English dialect identification. *Journal of Language and Social Psychology*, 18(1), 10-30.
- Racz, P. (2013). *Saliency in sociolinguistics*. Berlin: Mouton de Gruyter.
- Sailaja, P. (2009). *Indian English*. Edinburgh: Edinburgh University Press.
- Stephan, C. (1997). The unknown Englishes? Testing German students' ability to identify

- varieties of English In E. W. Schneider (Ed.), *English around the world 1* (pp. 93-108). Amsterdam: John Benjamins.
- Thomas, E. R. (2002). Sociophonetic applications of speech perception experiments. *American Speech*, 77(2), 115-147.
- Thomas, E. R., Lass, N. E. & Carpenter, J. (2010). Identification of African American speech. In D. R. Preston & N. Niedzielski (Eds.), *A reader in sociophonetics* (pp. 253-264). New York: Walter de Gruyter.
- Trakulkasemsuk, W. (2012). Thai English. In E. L. Low & A. Hashim (Eds.), *English in Southeast Asia* (pp. 101-112). Amsterdam: John Benjamins.
- Van Bezooijen, R. & Gooskens, C. (1999). Identification of language varieties: the contribution of different linguistic levels. *Journal of Language and Social Psychology*, 18(1), 31-48.
- Vieura, B., de Mareuil, P. B. & Adda-Decker, M. (2011). Characterisation and identification of non-native French accents. *Speech Communication*, 53(3), 292-310.
- Williams, A., Garrett, P. & Coupland, N. (1999). *Dialect recognition*. In D. R. Preston (Ed.), *Handbook of perceptual dialectology: volume 1* (pp. 345-358). Amsterdam: John Benjamins.
- Xu, W., Wang, Y. & Case, R. E. (2010). Chinese attitudes towards varieties of English: A pre-Olympic examination. *Language Awareness*, 19(4), 249-260.
- Yook, C. & Lindemann, S. (2013). The role of speaker identification in Korean university students' attitudes towards five varieties of English. *Journal of Multilingual and Multicultural Development*, 34(3), 279-296.