The role of variety recognition in Japanese university students’ attitudes towards English speech varieties

Robert M. McKenzie, University of Glasgow, Scotland

Language attitude studies have tended to assume that informants who listen to and evaluate speech stimuli are able to identify with consistent accuracy the varieties of English in question. However, misidentification could reduce the validity of any results obtained, particularly when it involves the evaluations of non-native English speaking informants, who are likely to have had less exposure to varieties of English speech. The present study investigated the perceptions of 558 Japanese university students of six varieties of English speech. The results indicated that whilst evaluations of speakers of UK and US English were particularly positive in terms of status, a Japanese speaker of heavily-accented English was rated most favourably in terms of social attractiveness. The findings from an additional dialect recognition question demonstrated that accurate identification had a significant positive effect upon the perceived status of native varieties of English, suggesting a tendency amongst the informants to look to native speakers to provide ‘notions of correctness’. The results also imply that Japanese learners retain representations of varieties of English speech and draw upon this resource, whether consciously or unconsciously, in order to identify and evaluate (speakers of) these speech varieties.

Introduction

Since Lambert et al. (1960) developed the matched-guise technique in order to investigate language attitudes in Canada, a high degree of consistency has been found from the data collected in the subsequent plethora of attitude studies conducted, allowing inferences to be drawn regarding the perceptions of native speakers of varieties of a language. It has been widely demonstrated, for instance, that standard speech varieties tend to be evaluated most positively by native speakers in terms of status/competence whereas non-standard varieties of English tend to be rated more highly in terms of solidarity/social attractiveness (Hiraga, 2005; Preston, 2004). This appears to be the case both when the judges are speakers of standard varieties and when the judges speak non-standard varieties of English.

In language attitude studies involving non-native speakers of English, researchers have found that English language learners, in a range of countries, are generally positive towards ‘the English language’, conceptualised as a single entity, e.g., Brazil (El-Dash and Busnardo, 2001), Singapore (Kwan-Terry, 1993) and Japan (Haarmann, 1989). However, research focusing on non-native perceptions of varieties of English is more limited. This is perhaps surprising given that the choice of pedagogical model(s) in English language classrooms should presumably take into account student attitudes towards particular varieties of the language.

In the case of Japan, however, there does exist a limited amount of research into attitudes towards varieties of English. For instance, Chiba et al. (1995) discovered that Japanese university students were more positive towards speakers of English in the inner
circle (where English is spoken as a native language for a substantial majority, such as the US and the UK) than speakers of either the outer circle (post-colonial countries, for instance, Malaysia and Sri Lanka) or the expanding circle (where English is learned as a foreign language, such as Japan). The results from a study conducted by Starks and Paltridge (1996) indicated that Japanese students preferred ‘American English’ (and to a lesser extent, ‘British English’) rather than ‘New Zealand English’ as learner goals. In a more recent study, McKenzie (2004) found that Japanese learners of English, studying in Scotland, were significantly more positive towards Scottish Standard English speech than Glasgow Vernacular speech. Cargile et al. (2006) investigated the perceptions of Japanese university students towards two varieties of US English: African-American Vernacular English (AAVE) and Mid-West United States English (MWUSE). The results of the study indicated that AAVE speakers were rated less positively in terms of status but more favourably in terms of social attractiveness, suggesting that the attitudes of the Japanese learners were broadly similar to those found throughout the USA.

Language Attitudes and Recognition

Despite the enormous amount of valuable research which has been undertaken in the field of psycholinguistics to understand the complex ways in which individuals perceive, process and encode spoken language (Clopper and Pisoni, 2005), language attitude researchers have tended to presume that informants who listen to and evaluate stimulus speech have accurately identified the varieties in question as socially or regionally recognised forms. There have, however, been recent calls to include a dialect recognition item in questionnaires, where participants are presented with voice samples and subsequently asked to rate them (e.g., Preston, 1999). Dialect recognition itself can be construed as the cognitive mapping of audible speech features on to the individual’s records of the usage norms of particular speech communities and to be achieved, the values of the variable features of the variety must be successfully identified and then appropriately mapped by the individual in question (Garrett et al., 2003: 208). Hence, although there is a counter-argument that the ability to recognise speech varieties may have no effect on the attitude of informants (e.g., they respond to the inherent value of the varieties in question), by the above account, when respondents are able to give a name to the variety under consideration, their evaluations are more likely to be based upon imposed social norms or connotations (Williams et al., 1999). Misidentification of speech varieties may, therefore, be a potentially confounding variable in language attitude studies and, as such, is liable to render the data more difficult to interpret. It should be noted, nevertheless, that patterns of misidentification, may also be useful in themselves. Speech varieties which have not been correctly identified, may, for instance, provide insights into the ideological framework of the respondents. Lindemann (2003), for example, maintains that listeners who are unable to correctly identify a particular speech variety may be likely to incorrectly identify the stimulus speech recording as a language or language variety with which they are more familiar. Lindemann (ibid.) believes that such identifications are frequently based on the ethnic associations of the listener, where, for example, a speaker from Canada may be wrongly identified as American, if indeed ‘Canada’ is not a particularly salient category for the listener.

A variety recognition question is, however, arguably of most value in attitude studies which involve evaluations by non-native speakers, who are likely to have had less
exposure to varieties of L2 speech than native speakers and, as such, may be less familiar with and have more difficulty in identifying particular varieties (i.e., they have more difficulty in achieving accurate cognitive mapping). Stephan (1997) maintains that although several attitude studies have attempted to measure the recognition rates of native speakers, not much is known about the ability of non-native learners to identify speakers’ place of origin solely from their speech. However, a study conducted by Ladegaard (1998), examining evaluations of English speech in Denmark, also attempted to measure recognition rates. Ladegaard concluded that ‘even though the judges are not native speakers of English, we may assume some degree of familiarity with the accents employed in this experiment since they sometimes appear in the media. It is possible therefore, that the subjects possess some kind of stored, “subconscious information”, based on previously acquired media-transmitted stereotypes’ (269). It would be profitable to conduct further research into the role of variety identification on language attitudes amongst learners of English in other linguistic and cultural contexts. The results obtained from such research are likely to validate (or not) the findings from the Danish study.

In light of the above discussion, the main objective of the present study was to ascertain how accurately and consistently Japanese learners of English could correctly identify specific varieties of English speech and to determine the influence (if any) that mis/identifications had on the learners’ attitudes towards varieties of English speech.

**Methodology**

**The varieties of speech selected**

In order to measure learners’ attitudes and recognition rates, six varieties of English speech were recorded. Since time constraints can unduly influence the validity of the data collected in speaker evaluation studies (Cargile and Giles, 2002), it was decided to give the informants a sufficient period of time to record their responses, as well as to present relatively lengthy speech samples. Thus, although it would have been interesting to present a greater number of varieties of English speech for evaluation, it was felt that listener-fatigue might compromise the validity of the data collected if more than six speech recordings of the required length were used. The varieties of English chosen consisted of four native (inner circle) varieties and two non-native (expanding circle) varieties. Two of the recorded native varieties of English are spoken in the UK: Glasgow vernacular (GV): and Glasgow Standard (GSE). The other two native varieties of English recorded are spoken in the United States: Southern United States English (SUSE) and Midwest United States English (MWUSE). The four native English speech varieties were selected specifically because native attitude research has indicated that together they constitute examples of the most and least favourably evaluated speech varieties in the UK (forms of Scottish Standard English, such as Glasgow Standard English, and Glasgow vernacular) (Macafee, 1994; Milroy, L., 1995) and in the US (Midwest US English and Southern US English) (Lippi-Green, 1997). Finally, recordings of two Japanese non-native speakers of English were included. Although both Japanese speakers were at an advanced level in English, one spoke moderately-accented Japanese English (MJJE) whilst the other spoke heavily-accented Japanese English (HJJE). The speakers were also selected for comparable voice qualities and overall, the authenticity of the recordings as representative samples of the varieties of English chosen were validated beforehand by
linguists from the USA (2) and Scotland (2) as well as by Japanese teachers of English (3). Nevertheless, it is important to bear in mind that each of the speech samples selected are merely an example of that particular variety and that other individuals in the same area or with the same social class, age or sex may not speak identically (Hiraga, 2005). In an attempt to minimise potential extraneous variables amongst the selected speech recordings, a number of factors were controlled. First, in terms of content, all speakers were recorded giving directions along a winding road on the same fictitious map. In addition, the recordings were screened for obvious references made to the speakers’ nationality, regional provenance or variety of English spoken and are broadly similar in length, ranging from 74 to 90 seconds. It was also decided to record only female speakers of English (although it is important to note that research suggests that male and female speakers of the same language variety may evoke different responses amongst listener-judges) (e.g., Kramarae, 1982). The age range of the speakers was relatively narrow; between 22 and 34 years of age, mean= 28, SD= 4.5.

The research instrument
The study consisted of two parts. First, to examine the informants’ evaluations of varieties of English speech, an indirect method of attitude measurement, the verbal-guise technique was employed. It is important when employing the VGT to select speakers very carefully to ensure that prosodic and paralinguistic features such as voice quality, speech rate and pitch remain broadly similar (for an overview see Garrett et al., 2003). There is evidence to suggest that different speech communities may react to any given adjective in different ways; in other words, reactions of informants are likely to be highly culture bound (El-Dash and Busnardo, 2001). Language attitude researchers should, thus, not suppose that traits will be salient for different populations. In the present study, a seven-point semantic-differential scale was specially constructed by means of a pilot study, where Japanese students, considered comparable to the informants selected for the main study, were asked to provide descriptions for each of the six speakers. In total, the eight most frequent descriptions (along with their bi-polar opposites) were selected and, in accordance with the majority of attitude studies involving rating scales in the field of social psychology, subsequently positioned in a randomised order to form the semantic-differential scale, i.e., the ‘socially most desirable’ traits were positioned sometimes on the left and sometimes on the right in order to avoid any left-right bias amongst the informants (for an overview see Oppenheim, 1992; Dornyei, 2003). Principal Components Analysis revealed that the traits selected reflected two non-overlapping dimensions of ‘competence’/’status’ (intelligent, confident, fluent, clear) and ‘social attractiveness’/’solidarity’ (gentle, pleasant, funny, modest), accounting for 26.49% and 15.14% of the variance.

Second, to ascertain identification (or not) of the six varieties of speech selected, the informants were asked the following two questions:

i) Where do you think the speaker comes from?

ii) How did you make this decision?

For the purposes of analysis, the identification was considered successful if the informants correctly recognised the country of the speaker (i.e., the USA, the UK or
Japan) and hence, the respondents were not required to identify the particular variety of English or region where it is spoken (if applicable).

The informants
The population selected for the study was principally Japanese nationals currently learning English at universities in Japan, either as a principal subject or as a major component within another discipline. Data was collected from 558 students of English, from eleven universities throughout Japan. The vast majority of the informants had completed a full six years of English education at school as well as a further one to three years at university. Although not central to the aims of the study, it is worth noting that a relatively large proportion of the informants (112) had spent three or more months abroad in ‘an English-speaking environment’. Moreover, although the great majority of students were undergraduates (513), a number of postgraduate students (45) also participated. The responses of a number of informants who did not report themselves as Japanese and/or as native speakers of Japanese were discarded. As a result, the sample appeared to be composed solely of university students of Japanese nationality, who spoke Japanese as a first language, were born in and, at the time of the data collection, lived in and studied in Japan. All data were complete with few exceptions. As the number of missing values was extremely small (seven in total) and seemingly random, a mean substitution strategy was employed. The age range of the sample was between 17 and 58, with the overwhelming majority of the respondents who participated in the study aged between 18 and 22 years of age (mean = 20.22, SD = 2.99).

Procedure
The data collection was undertaken in Japan over a two-month period, from October to December 2005. The informants listened to the speech samples two times, first for the verbal-guise study and again to complete the dialect recognition item. In the verbal-guise study, the order in which the speech samples were played was randomised. This decision was quite deliberate and undertaken to ensure any potential ordering effects were minimised. All other procedures involved in class visits were standardised.

Results and Discussion
Speaker evaluations: analysis of components extracted
Due to the random positioning of the positive/negative traits (see above), in the first stage of the analysis, a number of the informants’ responses were transposed, i.e., the lowest scores were converted to the highest scores and vice-versa (a value of 7 = most favourable rating and 1 = least favourable rating). The mean ratings (and standard deviations) for each speaker are summarised in Table 1. In order to investigate the significance of the means a one-way repeated measures analysis of variance (ANOVA) was conducted. The results demonstrated overall effects for competence: $F(5, 2785) = 266.90, p<0.05$; eta squared = 0.655, suggesting a large effect size and for social attractiveness: $F(5, 2785) = 57.09, p<0.05$; eta squared = 0.283, again indicating a large effect size. Since ANOVA does not indicate which sample means are producing the effect, follow-up pairwise
comparisons were also conducted (Bonferrini adjustments were made to control for Type I errors). The results of the pairwise comparisons are indicated in Table 1 below.

Table 1 Mean Evaluations (and Standard Deviations): Rankings for Competence and Social Attractiveness (N=558)

<table>
<thead>
<tr>
<th>Speaker</th>
<th>Competence</th>
<th>Social Attractiveness</th>
</tr>
</thead>
<tbody>
<tr>
<td>MWUSE</td>
<td>4.96</td>
<td>HJE</td>
</tr>
<tr>
<td></td>
<td>(1.03)</td>
<td>4.23 *</td>
</tr>
<tr>
<td>SUSE</td>
<td>4.95 *</td>
<td>GV</td>
</tr>
<tr>
<td></td>
<td>(0.98)</td>
<td>3.98</td>
</tr>
<tr>
<td>GV</td>
<td>4.29 *</td>
<td>SUSE</td>
</tr>
<tr>
<td></td>
<td>(1.00)</td>
<td>3.90</td>
</tr>
<tr>
<td>MJE</td>
<td>4.08 *</td>
<td>MJE</td>
</tr>
<tr>
<td></td>
<td>(1.11)</td>
<td>3.78</td>
</tr>
<tr>
<td>GSE</td>
<td>3.71 *</td>
<td>GSE</td>
</tr>
<tr>
<td></td>
<td>(0.92)</td>
<td>3.64 *</td>
</tr>
<tr>
<td>HJE</td>
<td>3.32</td>
<td>MWUSE</td>
</tr>
<tr>
<td></td>
<td>(0.97)</td>
<td>3.51</td>
</tr>
</tbody>
</table>

* indicates the position where there is a statistical difference (p<0.05) with the ratings below

GSE= Glasgow Standard English  MJE= Moderately-accented Japanese English
HJE= Heavily-accented Japanese English  MWUSE= Mid-West US English
SUSE= Southern US English  GV= Glasgow Vernacular

The results above demonstrate that in terms of competence, the Japanese informants’ rate speakers of native/inner circle Englishes significantly higher than speakers of non-native/expanding circle varieties. This finding parallels the results obtained from the limited number of studies previously conducted, which have measured perceptions of varieties of speech in Japan, where learners tended to express a preference for native varieties of English (see above). When the overall differences between the informants’ ratings are compared, a clear hierarchy emerges where speakers of US English are preferred, followed by the speakers of UK English with the Japanese speakers of English the least preferred. Moreover, it is interesting that although the two Japanese speakers were ranked lowest in terms of competence, the heavily-accented speaker was rated significantly less positively than the moderately-accented speaker. This result suggests that Japanese learners hold particularly negative perceptions of Japanese accented English, where the more recognisably ‘Japanese’ the speaker is perceived to sound, the more negatively she will be rated in terms of status.

In contrast, in terms of social attractiveness, the speaker of heavily-accented Japanese English was rated significantly more favourably than the other five speakers. The positive evaluation may indicate that the Japanese learners identify strongly with the speaker, i.e., there appears to be a high degree of solidarity with the HJE speech. It seems reasonable to assume that one reason for this is simply that the respondents are familiar with this variety. The speaker of moderately-accented Japanese English, however, was
rated much less positively on social attractiveness, and indeed this speaker was rated significantly less favourably than the speakers of non-standard/non-mainstream varieties of UK and US English. This relatively low rating may indicate that the MJE speaker is perceived as outgroup; i.e., although the speaker is afforded relatively high status, she may no longer be judged by the informants as a ‘true’ speaker of Japanese English. Furthermore, the rankings above indicate that when the evaluations for the social attractiveness of speakers of standard and non-standard varieties of inner circle English are compared, a preference is expressed for the non-standard speaker. This pattern is consistent with native speaker evaluations in the UK and in the US, where a preference for the non-standard variety on dimensions of social attractiveness also tends to be demonstrated (see above).

Identification (of speakers) of varieties of English

The next stage of the analysis was to ascertain how accurately and consistently the informants could identify the six varieties of English selected. The decision of whether an individual informant’s answers were considered correct or incorrect at times proved somewhat problematic, largely due to the idiosyncratic nature of the responses provided. For instance, although the first question specifically requested the listeners to identify the country where each of the speakers come from, several of the informants identified either the variety of English spoken or the nationality of the speaker. In addition, a relatively large number of respondents also frequently identified the provenance of the Scottish speakers (GSE and GV) as ‘the UK’ or ‘Britain’, which again made categorisation problematic. Under the circumstances (because the listeners were learners of English studying in Japan), a decision was taken not to impose an unrealistically narrow interpretation of the informants’ responses. For this reason, inaccuracies in both terminology and spelling were liberally interpreted and variations on both ‘the UK’ and ‘Britain’ were accepted as appropriate identifications of the provenance of the GSE and GV speakers. The percentages of the correctly and incorrectly identified place of origin for the six speakers are summarised below.

Table 2 Percentages (and Frequencies) of Correct and Incorrect Identifications for Country of Origin of Speaker (N= 558)

<table>
<thead>
<tr>
<th>Recognition</th>
<th>GSE (179)</th>
<th>HJE (503)</th>
<th>SUSE (330)</th>
<th>MJE (167)</th>
<th>MWUSE (305)</th>
<th>GV (173)</th>
</tr>
</thead>
<tbody>
<tr>
<td>Correct</td>
<td>32.08</td>
<td>90.14</td>
<td>59.14</td>
<td>29.93</td>
<td>54.66</td>
<td>31.00</td>
</tr>
<tr>
<td>Incorrect</td>
<td>67.92</td>
<td>9.86</td>
<td>40.86</td>
<td>70.1</td>
<td>45.34</td>
<td>69.00</td>
</tr>
</tbody>
</table>

GSE= Glasgow Standard English  MJE= Moderately-accented Japanese English
HJE= Heavily-accented Japanese English  MWUSE= Mid-West US English
SUSE= Southern US English  GV= Glasgow Vernacular
The recognition rates detailed above may appear somewhat low when compared to the results of previous variety identification studies, involving either native speakers or non-native speakers, where higher rates of recognition were found (e.g., Purnell et al., 1999; Stephan, 1997). In contrast, the vast majority of these studies asked listeners to select from a predetermined list, thus, limiting the types of misidentification possible. However, in the present study, as the recognition questions were open-ended (and hence, no predetermined list of response options was provided), it would be reasonable to expect the listeners’ recognition rates to be lower. Moreover, because the choice of the speakers’ place of origin was not limited in any way, the patterns of misidentification found amongst the listeners’ responses to the open-ended recognition questions also have the potential to provide greater understanding of the ideological framework of the informants (see above). For these reasons it was considered worthwhile to ascertain the listeners’ correct and incorrect rates of categorisation of ‘native English speech’ (in the case of the UK and US speakers) and of ‘non-native English speech’ (in the case of the Japanese speakers). The findings are presented below.

Table 3 Percentages (and Frequencies) of Correct and Incorrect Categorisation for Native/Non-Native Speaker (N= 558)

<table>
<thead>
<tr>
<th>Recognition</th>
<th>Speaker</th>
<th>GSE</th>
<th>HJE</th>
<th>SUSE</th>
<th>MJE</th>
<th>MWUSE</th>
<th>GV</th>
</tr>
</thead>
<tbody>
<tr>
<td>Correct</td>
<td></td>
<td>60.76 (339)</td>
<td>93.37 (521)</td>
<td>82.97 (463)</td>
<td>53.86 (301)</td>
<td>82.62 (461)</td>
<td>46.41 (259)</td>
</tr>
<tr>
<td>Incorrect</td>
<td></td>
<td>39.24 (219)</td>
<td>6.63 (37)</td>
<td>17.03 (95)</td>
<td>46.14 (257)</td>
<td>17.38 (97)</td>
<td>53.59 (299)</td>
</tr>
</tbody>
</table>

The data presented in Table 2 and Table 3 above demonstrated that there were great differences between the informants’ recognition rates for the place of origin of the six speakers. The country-specific recognition rates for both the mainstream variety of US English, MWUSE (54.66%), and the non-mainstream variety of US English, SUSE (59.14%) were relatively high. The most plausible explanation for the high ‘hit rate’ for the two varieties is the prevalence of American culture in Japanese society. This prevalence is demonstrated by the dominance of US news and movies in the English language media in Japan (Stanlaw, 2004) and hence, a current tendency for reliance upon US varieties of English to provide the models and norms for English language use in the country (Kubota, 1998). It is interesting to note that there were high levels of identification of the SUSE (82.97%) and MWUSE speech (82.62%) as inner circle varieties of English. Many of the learners, in response to the question ‘how did you make this decision?’ typically commented upon the ‘clarity’ of the speech and their ‘ease of understanding’, which again suggests a relative familiarity with US varieties of English. (The possibility exists, however, that the higher recognition rate found for the (non-standard) SUSE speaker may be explained by the comprehensive coverage of the events of Hurricane Katrina and subsequent prevalence of speakers of Southern United States English in the US dominated English language media in Japan in the months prior to the
fieldwork visit, and hence, the greater awareness of these varieties that learners of English may have had during the data collection period).

However, the informants demonstrated considerably more difficulty in terms of the correct identification of the standard variety of UK English, Glasgow Standard English (32.08%), and the non-standard variety, Glasgow vernacular (31.00%), selected for evaluation. The most plausible explanation for the relatively low hit-rates is that, because of a lack of exposure in Japan, the learners are generally unfamiliar with localised UK varieties of English speech and thus do not have sufficient experience and awareness of these forms of speech to achieve accurate identification. Furthermore, the higher rate of recognition for the Glasgow Standard English speech in comparison with the Glasgow vernacular speech is likely to reflect the somewhat greater exposure afforded to standard varieties than non-standard varieties of UK English in the English language media and in the language classroom in Japan. On the other hand, a comparatively high proportion of learners (60.76%) recognised the GSE speech as a native variety, which strongly suggests an ability to distinguish between native and non-native varieties of English. Such awareness is manifested in many of the learners’ comments, where there was a propensity to describe the ‘fluency’ of the GSE speaker and the ‘distinctiveness’ of the speech from ‘American English’. In contrast, a much lower proportion of the informants were able to recognise the non-standard variety of UK English, GV (46.41%), as inner circle speech. Intriguingly, a relatively large proportion of the informants (26.88%) perceived the place of origin of the GV speaker as ‘other Europe’ (i.e., from the expanding circle), and in particular, from France, Germany or Italy. Whilst this finding clearly demonstrates a lack of exposure to local varieties of non-standard UK speech in Japan, it also suggests that an inherent linguistic feature(s) of the Glasgow vernacular itself may have played an important role in the informants’ categorisation; those learners who failed to recognise the provenance of the GV speaker and identified her as French, German or Italian frequently commented upon the speakers pronunciation of the phoneme /r/, the implication of which, is that it is this specific linguistic feature of Glasgow vernacular speech which triggered misidentification amongst these particular informants. It should also be noted, nevertheless, as in the current study, that ‘the use of natural speech makes it more difficult to isolate the precise linguistic variants that naïve listeners attend to in making explicit categorisation judgements…further research using both natural and synthetic stimuli is needed to explore the role of individual linguistic variants, and the combinations of variants, that are salient for naïve listeners in perceptual dialect categorization studies’ (Clopper and Pisoni, 2006: 214). This is particularly the case with research involving language learners.

The HJE speaker was the most accurately identified (90.14%) of the six speakers, which clearly demonstrates a high degree of familiarity with heavily-accented Japanese English speech. Moreover, the generally positive ratings for the HJE speaker on the dimension of social attractiveness (see above) suggests that there exists a high degree of solidarity with the speaker amongst the learners, many of whom are themselves likely to be speakers of heavily-accented Japanese English. Nevertheless, the relatively low ratings for the competence of the HJE speaker demonstrated that the learners generally perceive heavily-accented Japanese English as both ‘lacking in prestige’ and ‘incorrect’. Further evidence of ambiguity in the informants’ attitudes towards the HJE speaker (and hence,
heavily-accented Japanese English) is reinforced by the responses of the learners to the question ‘how did you make this decision?’; where, on the one hand, there was a tendency to comment upon the ‘ease of comprehensibility’ and ‘familiarity’ of the speech and, on the other, the ‘lack of fluency’ and ‘incorrect pronunciation’ of the speaker. The recognition rate for the place of origin of the MJE speaker (29.93%), in contrast, was very much lower. It is indeed possible that the relatively low level of accurate identification is solely as a result of the impact of prolonged periods of academic study in the UK and the USA on the spoken English of this particular speaker. However, it is interesting to note that many more informants perceived the MJE speaker to be from the expanding circle (62.37%) than from the inner circle of English use (24.90%). Therefore, the ability to distinguish between expanding circle and inner circle varieties of English that the learners appear to possess again demonstrates that the native/non-native distinction is paramount for the informants in the identification process, and suggests that recognition is occurring at some level of awareness, (despite a general tendency to actively categorise the speaker of moderately-accented Japanese English as outgroup, in particular, as other Asian (13.98%) or European (14.87%) countries in the expanding circle.

In the majority of previous language attitude studies which included a dialect recognition item, recognition of a variety was generally construed as a process of cognitive mapping (see above). However, in the case of the present study, the association between high levels of identification and solidarity with the HJE speaker and, in the case of the MJE speaker, a low hit-rate and a lack of solidarity, indicates that processes such as *claiming* (for identification) and *denial* (for misidentification) (see Eagly and Chaiken, 1993; Devine, 1995) may be important for the learners’ recognition of varieties of English spoken by Japanese. The existence of such processes raises the possibility that Japanese learners’ recognition of forms of English spoken by Japanese speakers of English are influenced by ‘active in-grouping processes’ (Tajfel, 1974), where (as in a study by Garrett *et al.* of native speaker attitudes towards varieties of English in Wales), variety recognition may be ‘part of more elaborate process of “social cognition”, reflecting ideologies and preferences in listeners’ communities and strategies in representing them’ (Garrett *et al.*, 2003: 227).

Williams *et al.* (1999), in a study of Welsh teenagers’ attitudes towards (speakers of) English in Wales, go a step further, arguing that *affect* (i.e., emotions, moods and preferences) may also play a role in dialect recognition. Williams *et al.* found that the teenagers did not only recognise (or fail to recognise) speakers as belonging to specific communities, but also tended to appropriate a ‘likeable speaker’ into their own ingroup. Williams *et al.* concluded that there might be a group-level affective dimension of variety recognition which is ‘likely to dominate in recognition tasks in which accurate cognitive mapping cannot be achieved: for example, when listeners are inexperienced’ (358).

Because language learners are likely to have had less exposure to varieties of English speech than native speakers of the language (i.e., they are comparatively less experienced), the claim by Williams *et al.* may have a particular relevance to the recognition of varieties of the L2. Although compelling, the claim is highly speculative, because, at present, no convincing theory exists which can account for the role of emotion in dialect recognition. Moreover, although psychologists, for analytic convenience, tend to divide affect and cognition, and there is some evidence that people
can know how they feel about an object before they recognize it, e.g., when listening to opening bars of music (Fiske and Taylor, 1991), in the course of any given individual’s lived experience, affect and cognition occur in a ‘simultaneous mix’ (ibid: 410). There are also problems in comparing affect and cognition, since in different disciplines they have been distinguished in different ways, e.g., sensory vs. inferential, physiological vs. mental, motor vs. perceptual, innate vs. learned, preference vs. knowledge and liking vs. discrimination (ibid: 457). Nevertheless, the focus on emotion reflects a current trend in social psychology generally, where researchers, who have traditionally focussed only on describing the cognitive processes and structures which influence social behaviour, are currently also turning their attention to the role that affect may also play (e.g., Eich and Schooler, 2000; Forgas, 2001). As a result, the study of the interaction between affect and cognition is currently one of the most active and rapidly developing areas within psychological science. Indeed, researchers have already demonstrated, for instance, that ‘racial schemas have a strong affective component, so that the mere sight of an individual from a particular group may trigger emotions like fear and suspicion and evaluative judgements which are negative and derogatory’ (Augoustinos and Walker, 1995: 48). Hence, given social psychologists ‘increased knowledge of affective influences on individual-level judgements and processing of information’ (Kelly, 2001: 177), it would be of potential worth, if a suitable methodology can be developed, to conduct further research into the interaction of affect and cognition in dialect recognition when conducting attitude studies involving the evaluations of non-native speakers of English, especially when speakers from the listener-judges own country are selected to be the object of evaluation.

The patterns of misidentification are also interesting from another point of view as it is reasonable to assume that if learners had little or no experience and awareness of the varieties which they did not identify accurately, their responses would be random (Williams et al., 1999). However, a high degree of consistency, in fact, was found amongst the informants’ misidentifications for the place of origin for all six speakers. For instance, as described above, informants who failed to identify the provenance of the MWUSE and SUSE speakers as ‘the United States’ or the GSE speaker as ‘Scotland/the UK’ were, nevertheless, generally able to recognise the speech as inner circle English. This finding supports the assertion that the native/non-native distinction is a salient one for the informants and again strongly suggests that there is a tendency for the Japanese learners to classify speakers initially as either native or non-native before attempting to further categorise them; perhaps based upon more specific ethnic associations (see above).

**Effects of Mis/identification**

The final section of the study investigated whether any differences found between correct and incorrect identifications in the dialect recognition section of the research instrument affected the mean evaluations of each of the six speakers in terms of competence and social attractiveness. The first stage of the analyses was to calculate descriptive statistics for the competence and social attractiveness of all six speakers according to correct and incorrect identifications. A mean value of seven corresponds to the most favourable rating and, in contrast, a value of one indicates the least favourable rating. This data is summarised below:
Table 4 Mean Evaluations (and Standard Deviations) for Competence and Social Attractiveness according to Correct and Incorrect Identifications (N=588)

<table>
<thead>
<tr>
<th>Speaker</th>
<th>Competence</th>
<th>Social Attractiveness</th>
</tr>
</thead>
<tbody>
<tr>
<td></td>
<td>Correct</td>
<td>Incorrect</td>
</tr>
<tr>
<td>GSE</td>
<td>4.22 (1.16)</td>
<td>4.02 (1.08)</td>
</tr>
<tr>
<td>HJE</td>
<td>3.30 (0.96)</td>
<td>3.49 (0.96)</td>
</tr>
<tr>
<td>SUSE</td>
<td>5.03 (0.96)</td>
<td>4.82 (1.01)</td>
</tr>
<tr>
<td>MJE</td>
<td>3.75 (0.92)</td>
<td>3.69 (0.92)</td>
</tr>
<tr>
<td>MWUSE</td>
<td>5.10 (1.00)</td>
<td>4.79 (1.04)</td>
</tr>
<tr>
<td>GV</td>
<td>5.10 (1.00)</td>
<td>4.29 (1.00)</td>
</tr>
</tbody>
</table>

The results from Table 4 demonstrate a general tendency towards more positive evaluations when the place of origin for the speakers of inner circle English are correctly identified. In order to determine the significance of the effects of mis/identification, a series of one-way between groups multivariate analyses of variance (MANOVA) were subsequently conducted. The results indicated that although mis/identification had some bearing on the informants’ ratings for social attractiveness, in each case the differences between the mean ratings failed to reach significance. However, in terms of competence, significant effects were found for the evaluations of three speakers: SUSE F(1, 556)= 5.460, p<0.05 (p= 0.02); MWUSE F(1, 556)= 13.104, p<0.05 (p= 0.000); GV F(1, 556)= 24.357, p<0.001 (p= 0.000). It is important to stress that in each case recognition of the speaker’s place of origin resulted in a significantly more positive evaluation.

These findings may indicate that, as far as ratings of inner circle varieties of English are concerned, recognition has a positive effect on perceptions of the competence of the speakers of these varieties, and hence, on the prestige of inner circle varieties of English speech. In turn, the results imply that, as informants who recognised a particular variety of inner circle English were most likely to be familiar with it, it may be that familiarity itself also had a positive influence on the learners’ attitudes towards the status of native varieties of English speech.

As described above, a plethora of language attitude studies have demonstrated that native speakers of English consistently evaluate standard varieties of inner circle English more highly in terms of prestige than non-standard varieties. J. Milroy (1999) has attributed the consistency found in these studies to the existence of a ‘standard language ideology’, often promoted indirectly by linguists, where in any given geographical area, a
specific variety of English is recognised as ‘the standard’. This variety is thus considered to embody ‘notions of correctness’ and, as such, speakers of this ‘prestige variety’ are afforded a degree of respect in the society as a whole (Bex and Watts, 1999), although they may be downgraded in terms of social attractiveness (solidarity). In the case of the present study, the significantly more favourable ratings for the competence of speakers of varieties of inner circle English whose provenance was identified points to the construction of a ‘native speaker ideology’ amongst the Japanese informants and implies that these learners of English tend to look towards (both standard and non-standard) varieties of inner circle English for notions of correctness. This is broadly compatible with the view of Tsuda (1997) who maintains that perceptions of English as an international language in Japan has resulted in the glorification of speakers of varieties of inner circle English, a process he defines as ‘Anglomania’.

**Conclusion**

The discussion above demonstrates that what constitutes ‘recognition’ of a language or a language variety is a complex process. Nevertheless, it has been widely demonstrated by speech perception researchers that, through a combination of experience with and exposure to both the speech community and the world in general, individuals retain a memory of the varieties of their native language(s) to the extent that they can imitate, identify the place of origin and make judgements about social characteristics of speakers of these varieties (Clopper and Pisoni, 2005). As described previously, Ladegaard (1998) found that non-native informants in Denmark are capable of making comparable discriminations between varieties of the English, and likely to base their judgements on stored, subconscious information. The findings of the present study imply that the Japanese learners also retain representations of varieties of English and drew upon this resource, whether consciously or unconsciously, in order to both complete the recognition task and to assign individual characteristics to the speakers in the verbal-guise section of the study. Nevertheless, in order to validate the findings, there is a clear requirement for further research of a similar nature to be conducted involving English language learners studying in a range of countries and the use of different methodological approaches, for instance, by incorporating techniques from the field of perceptual dialectology (Preston, 1999). Nevertheless, it is hoped that the findings from the present study demonstrate the value of including a dialect recognition item in the research instrument when measuring the perceptions of non-native learners of varieties of English speech generally, as well as providing a basis for comparison with future studies, where the specific objective is to measure the recognition rates and patterns of categorisation of varieties of English speech amongst Japanese learners of English.

**Acknowledgements**

The financial aid provided by the Gen Foundation, London is gratefully acknowledged.
References

Ladegaard, H. (1998) National stereotypes and language attitudes: the perception of
British, American and Australian language and culture in Denmark. *Language and Communication* 18, 251-274


