

The assessment of foreign accent by native and non-native judges

Francisco Gallardo del Puerto, University of the Basque Country
Esther Gómez Lacabex, University of the Basque Country
María Luisa García Lecumberri, University of the Basque Country

1. INTRODUCTION

The term 'foreign accent' (FA) is used to refer to the pronunciation of a language that shows deviations from native norms. These deviations occur at the phonetic and phonemic level, and at the segmental and prosodic levels, characterising the user of a second language (L2) as a non-native speaker. There is a consensus that the pronunciation of an L2 is deeply affected by the native language, above all in the early stages of acquisition (Ellis, 1994). In fact, one can easily detect the origin of some L2 speakers just by listening to the way the L2 is pronounced by them. Nevertheless, there are also reports of learners who achieve near-native pronunciations (Bongaerts, 1999).

FA is usually assessed holistically with the help of judges. In an attempt to discover whether FA affects speech comprehensibility and acceptability, native speakers of the language in question are required to consider the seriousness of learners' deviations from native phonological norms on the basis of such scales as degree of FA, intelligibility or irritation. Previous research findings (Fayer & Krasinski, 1987; Munro & Derwing, 1995; García Lecumberri & Gallardo del Puerto, 2003) indicate that a lower degree of FA is associated with higher intelligibility and lower irritation, and that degree of FA is the most severely judged measure.

However, the assessment of how foreign, intelligible or irritating an accent is also depends on the raters' own linguistic background. Some studies (Calloway, 1980; Flege, 1984; Thompson, 1991) have shown that linguistically experienced listeners are more reliable than inexperienced judges in estimating L2 learners' speech intelligibility, irritation and/or accentedness. Baetens Beardsmore (1979) verified that listeners' type and degree of bilingualism/multilingualism affected the degree of perceived FA and, more particularly, its intelligibility and acceptability. In the same vein, other researchers (Flege, 1988; Scovel, 1977) discovered that non-native judges' ability to recognise FA appears to increase proportionally with command and experience with the foreign language (FL). In fact, some experiments insist on the similarity between native and non-native judges when assessing FA, as evidenced by the lack of statistical differences between the FA judgements by these two types of listener (Flege, 1988; MacKay et al., 2006).

Non-native instructors are frequently in charge of teaching L2s in formal instruction settings. Consequently, it is often the case that non-natives are the ones who assess learners' pronunciation development. The motivation of the present study was to ascertain the reliability and differences between two types of judges: naïve native judges vs. linguistically trained non-native judges.

2. METHODOLOGY

28 Basque-Spanish bilingual speakers were analysed. They had been studying English as a FL at school for 6.5 to 7.5 years (mean 6.8) from age 8 and had received no extra-curricular exposure to English. By the time they took the test, their ages ranged from 14 to 16 (mean: 14.6) and their school exposure to English from 693 to 1155 hours (mean: 850.6). The instrument employed to analyse learners' FA in English was a wordless picture story-telling activity ('*Frog, where are you?*' by Mayer, 1969). Learners' productions were recorded and two-minute excerpts were subsequently randomized and judged by ten listeners. Three different nine-point scales were utilized in order to assess students' pronunciations: degree of FA (1 = very strong; 9 = very slight), FA intelligibility (1 = impossible to understand; 9 = extremely easy to understand), and FA irritation (1 = extremely annoying; 9 = not at all annoying). There were two groups of judges: native judges (NJs) and non-native judges (NNJs). The NJs were 5 monolingual speakers of British English, post-graduates in a Computer Science department in the UK, who had little linguistic/phonetic training. The NNJs were 5 native speakers of Spanish, academics or postgraduates in an English Language department in Spain, who were proficient in English, experienced as teachers of English and phonetically/linguistically trained.

3. RESULTS

In order to compare the judgements made by NJs and NNJs, two types of analyses were carried out: correlations and analyses of variance. Correlation analyses were firstly conducted in order to verify the degree of consistency among three FA measures (DFA = degree of FA, FAIN = FA intelligibility, and FAIR = FA irritation) for the two judge groups (see table 1). It was observed that the highest values corresponded to the correlations between FA intelligibility and irritation and that NJs' correlations were always lower than NNJs', especially when DFA was involved in the comparison. In other words, NNJs exhibited a higher inter-measure consistency than NJs as all their FA measurements were highly correlated (Pearson indexes over .80). For NJs, however, degree of FA seemed to be less strongly connected to FA intelligibility and irritation (Pearson indexes around .70) than for NNJs.

	<i>DFA - FAIN</i>		<i>DFA - FAIR</i>		<i>FAIN - FAIR</i>	
	<i>Pearson Index</i>	<i>Significance</i>	<i>Pearson Index</i>	<i>Significance</i>	<i>Pearson Index</i>	<i>Significance</i>
<i>NJ</i>	.705	<.001	.680	<.001	.904	<.001
<i>NNJ</i>	.845	<.001	.868	<.001	.932	<.001

Table 1: inter-measure correlations by NJs and NNJs

Additionally, correlations between NJs and NNJs for the three FA variables were conducted. As shown in table 2, NJs and NNJs turned out to be highly correlated for FAIR, moderately correlated for FAIN, and slightly correlated for DFA. That is to say, NJs and NNJs showed a higher correlation in their assessment of FA irritation than in their rating of FA intelligibility and degree of FA.

<i>DFA</i>		<i>FAIN</i>		<i>FAIR</i>	
<i>Pearson Index</i>	<i>Significance</i>	<i>Pearson Index</i>	<i>Significance</i>	<i>Pearson Index</i>	<i>Significance</i>
.583	.001	.713	<.001	.837	<.001

Table 2: NJ-NNJ correlations

In order to compare more specifically the judgements made by NJs and NNJs, a comparison of the means, standard deviations and ranges obtained for each FA scale (min = 1; max = 9) by the two types of judge is displayed in table 3. Higher scores in each of the scales are always related to more positive assessments about the learners' accent, that is, to lower degree of FA, to less irritating FA, and to less unintelligible FA.

	<i>NJs</i>		<i>NNJs</i>		<i>T</i>	<i>Sig.</i>
	<i>Mean (Range)</i>	<i>S. D.</i>	<i>Mean (Range)</i>	<i>S. D.</i>		
<i>DFA</i>	3.16 (2-6)	.94	3.17 (1-6)	1.03	-.083	.934
<i>FAIN</i>	4.13 (2-7)	1.64	5.50 (3-8)	1.60	-5.894	.001
<i>FAIR</i>	4.61 (2-7)	1.14	4.69 (2-8)	1.73	-.379	.708

Table 3: NJ-NNJ comparisons for DFA, FAIN and FAIR

Table 3 results indicated that DFA mean scores exhibited by the two groups of judges were very similar. Unsurprisingly, the T-test yielded no statistical significance for DFA. Thus, NJs and NNJs estimated the degree of FA exhibited by our learners as a group as very similar. Similar results were also observed for FAIR, indicating that the productions of our participants as a group were considered to be equally irritating by NJs and NNJs. As far as DFA and FAIR are concerned, it is worth noting that NNJs used a slightly wider range for their scores and that their standard deviations were slightly wider too. Nevertheless, intelligibility assessment showed some differences between NJs and NNJs, as the mean scores shown by NNJs turned out to be higher than those of NJs, differences being statistically significant. In other words, not surprisingly, NNJs judged learners' productions as considerably more intelligible than NJs. Besides, NJs and NNJs displayed similar ranges and standard deviations when evaluating intelligibility.

4. DISCUSSION

Results show that phonetically and/or linguistically trained non-native listeners considered the three different measures of FA to be more interrelated than untrained native listeners, as shown by the stronger correlations among degree of FA, FA intelligibility and FA irritation amongst the former group. Specifically, for NNJs a stronger FA was more connected to less intelligibility and was more irritating than for NJs. This could be interpreted as trained judges being more focused when assessing how degree of FA affects speech intelligibility and irritation, which would be in line with previous findings that trained listeners are more reliable in estimating the foreign-ness of accents (Calloway, 1980; Flege, 1984; Thompson, 1991). Untrained listeners, however, might rely more strongly on factors other than phonological/phonetic cues when assessing the intelligibility and the irritation of learners' productions, such as grammatical accuracy (Varonis & Gass, 1981) or fluency (Anderson-Hsieh & Koehler, 1988). It may even be the case that DFA scores themselves are also influenced by non-phonetic variables (García Lecumberri & Gallardo del Puerto, 2003; Munro & Derwing, 1995; Varonis & Gass, 1981). This interpretation would be in agreement with our findings that DFA is the measure that least correlates NJ and NNJs.

Results also confirm previous findings (Fayer & Krasinski, 1987; García Lecumberri & Gallardo del Puerto, 2003; Munro & Derwing, 1995) in that a lower degree of FA is associated with higher intelligibility and lower irritation and in that accentedness judgements are harsher than intelligibility or irritation measures. However, our NJs and NNJs differed in how intelligible learners' accents were. NNJs judged learners' accents more intelligible than NJs. It is worth remembering that NNJs were native speakers of

Spanish and/or Basque and had worked as teachers of English to Spanish/Basque natives. Their experience with both Spanish and Basque phonological systems and their familiarity with Spanish/Basque-accented English aided their understanding of learners' productions. Additionally, their higher ability to understand participants' speech could also be ascribed to the fact that learners' productions sometimes contained Spanish and/or Basque code-switching, a fact which probably affected NJs' comprehension negatively. However, as indicated by the lack of statistical differences between NJs and NNJs in the other two FA scales, NNJs' higher understanding did not lead them to consider lower degrees of FA and FA irritation, as compared to NJs. That is, NNJs were equally able to judge the accentedness and irritation of FA by disregarding the fact that they could understand the students better than NJs. Consequently, our results would support those studies that have indicated that NNJs are as capable as NJs in detecting accentedness (Flege, 1988; MacKay et al., 2006).

REFERENCES

- Anderson-Hsieh, J. & Koehler, K. (1988) The effect of foreign accent and speaking rate on native speaker comprehension. *Language Learning*, vol. 38/4, pp. 561-613.
- Baetens Beardsmore, H. (1979) The recognition and tolerance level of bilingual speech. *Working Papers on Bilingualism*, vol. 19, pp. 116-128.
- Bongaerts, T. (1999) Ultimate attainment in L2 pronunciation: the case of very advanced late L2 learners. In Birdsong, D. (ed.) *Second Language Acquisition and the Critical Period Hypothesis* (pp. 133-160). New Jersey: Lawrence Erlbaum.
- Calloway, D.R. (1980) Accent and the evaluation of ESL oral proficiency. In Oller Jr., J.W. & Perkins, K. (eds.). *Research in Language Testing* (pp. 102-115). Rowley, Massachusetts: Newbury House.
- Ellis, R. (1994) *The study of second language acquisition*. Oxford: Oxford University Press.
- Flege, J.E. (1984) The detection of French accent by American listeners. *Journal of the Acoustical Society of America*, vol. 76, pp. 692-707.
- Flege, J.E. (1988) Factors affecting degree of perceived foreign accent in English sentences. *Journal of the Acoustical Society of America*, vol. 84, pp. 70-79.
- García Lecumberri, M.L. & Gallardo del Puerto, F. (2003) English FL sounds in school learners of different ages. In García Mayo, M.P. & García Lecumberri, M.L. (eds.) *Age and the acquisition of English as a foreign language* (pp. 115-135). Clevedon: Multilingual Matters.
- Fayer, J.M. & Krasinski, E. (1987) Native and non-native judgements of intelligibility and irritation. *Language Learning*, vol. 37/3, pp. 313-327.
- MacKay, I.R.A., Flege, J.E. & Imai, S. (2006) Evaluating the effects of chronological age and sentence duration on degree of perceived foreign accent. *Applied Psycholinguistics*, vol. 27, pp. 157-183.
- Mayer, M. (1969) *Frog, Where Are You?*. New York: Dial Press.
- Munro, M.J. & Derwing, T. M. (1995) Foreign Accent, comprehensibility, and intelligibility in the speech of second language learners. *Language Learning*, vol. 45, pp. 73-97.
- Scovel, T. (1977) The ontogeny of the ability to recognize foreign accents. In Henning, C.C. (ed.) *Proceedings of the Second Language Research Forum* (pp. 38-47). New York: UCLA.
- Thompson, Y. (1991) Foreign accents revisited: The English pronunciation of Russian immigrants. *Language Learning*, vol. 41/2, pp. 177-204.
- Varonis, E.M. & Gass, S. (1981) The comprehensibility of non-native speech. *Studies in Second Language Acquisition*, vol. 4, pp. 114-136.