

A COMPARISON OF THE ABILITY OF L1 AND L2
SPEAKERS TO RANK TYPES OF OVERT DEVIANCE
WITHIN ENGLISH SENTENCES.

J.R. Weatherhead.

University of Papua New Guinea

This paper reports a continuation of work started at the University of Lancaster in 1973. At Lancaster I asked native speakers of English to rank sentences containing errors in order of their relative acceptability. The effects of context, order of presentation, presence or absence of a correct "criterion" sentence, and different lexical items were investigated; and an attempt was made to assess differences between language teachers and university students, and native speakers of two different dialect groups.

Two aspects of the Lancaster work - the degree of homogeneity of judgements displayed by these native speakers, and the manner in which judgements of language teachers differed from those of native speakers - are of importance to this study and will be discussed below.

In this study I gave three groups of second-language speakers, all Papua Guineans, the same task as the Lancaster groups. These L2 groups consisted of (1) twenty-two first year students from UPNG; (2) thirty-one form 4 students from Iarowari High School; and (3) twenty-nine form 1 students also from Iarowari High School. The object of the experiment was to see whether L2 speakers at these levels of competence can make judgements of relative acceptability when presented with sentences which are deviant in various ways.

Five sentences (see Appendix) were presented in all possible combinations of pairs, making ten pairs of sentences in all. The subjects were asked to decide which of the two sentences was in the "best" English, and these judgements of acceptability were collated into a rank order for the five sentences. There were, inevitably, a number of circular judgements, where for example sentence A was judged better than sentence B, which was judged better than sentence D although D was judged better than A. In this case (there were 8 examples of this type) sentences A, B and D would be assigned equal ranks. These rank orders were then used in the statistical treatment.

The basic statistic calculated was Kendall's concordance coefficient, usually labelled W, which gives "a measure of the overall correlation when there are k sets of rankings of the same n objects. Put in another way it provides a measure of the agreement between the rankings of n objects by k judges. This coefficient is defined as

$$W = \frac{12 \sum s^2}{k^2 (n^3 - n)}$$

where S is the sum of the squares of the deviations of the total of the ranks obtained by each object from the average of these totals "(Maxwell, 1961: 119)

The significance of values of W can be found fairly easily by applying a correction for continuity and obtaining a variance ratio which can then be referred to the F-distribution for the appropriate degrees of freedom. I was interested, however, not merely in the values of W obtained but in differences between those values; and here, unfortunately, no test of significance has been devised.

Table 1 shows the values of \underline{W} obtained for the various groups - at Lancaster, and in Papua New Guinea.

Table 1

Values of \underline{W} obtained by various groups

** = significant at 1%; * = significant at 10%

	<u>group</u>	<u>W</u>	<u>K</u>
	L1 speakers	} 0.53 **	20
	(dialect groups)		} 0.41 **
L1	University students	0.44 **	20
L1	Language teachers	0.44 **	20
L2	UPNG students	0.28 *	22
L2	Form 4 students	0.15	31
L2	Form 1 students	0.20	29

It can be seen that the values for \underline{W} are much higher for the L1 speakers than for the L2 ones. This might have been expected, but since there exists no test for significance of differences in \underline{W} it remains possible that the differences are due to chance. What can be said, however, is that in all cases the L1 speakers are making judgements in a fairly homogenous (the \underline{W} - correlation - values are high) and non-random fashion compared to the L2 speakers. Of the groups of L2 speakers only that of UPNG students is significantly likely to be behaving in a non-random fashion; and that at only the 10% level rather than the 1% of the L1 speakers.

It seems, then, on this (obviously inconclusive) evidence, that Papua New Guinea High School students are less able to make judgements of relative acceptability when presented with sentences which are deviant in some way from normal English usage. L2 UPNG students, and L1 speakers, are more able to make such judgements.

The next obvious question is whether the judgements of L1 and L2 speakers are the same. For this the mean rank assigned to each sentence by each group was calculated. Differences in rank within groups can be assessed for significance, using the t - distribution (an article by Stuart (1951) gives the method).

Table 2 gives the mean rank order of the five sentences for the different groups. The letters A, B, C, D, and E designate the sentence, brackets indicate there is no significant difference in ranks assigned.

Table 2

Group mean rank orders assigned

		{	C	(AB)	E	D
L1	dialect groups	{	C	(AB)	E	D
L1	university students		C	(AB)	E	D
L1	language teachers		C	(ABE)		D
L2	UPNG students		C	(ABE)		D

It will be seen that the only difference in rank order is the location of sentence E. The three groups of naïve L1 speakers all put sentence E separate from A and B, the L1 language teachers and L2 UPNG students put A, B and E together; (the language teachers, incidentally, had all had experience in L2 areas).

Sentence C is a correct "criterion" sentence and sentence E differs from it by the replacement of a lexical item - "better" - with a less appropriate one - "lighter" (the sentences are given in the appendix below). A and B are deviant by virtue of a morphophonemic change ("feels" for "felt") and a change in phrase-structure rule (resulting in "feeling the better" for "feeling better:") respectively. D, to complete the picture,

involves a change in embedding process ("instead he felt better" for "instead of feeling better").

What is happening, therefore, is that L2 speakers and L1 language teachers are alike in judging an error in use of a lexical item as no worse than errors in phrase-structure and morphophonemics. The naive L1 speakers judge the lexical error to be worse than the other two.

An explanation of this for the L2 speakers is not hard to find. The methods of language teaching to which they have been exposed rely heavily on structural drills, with substitution of lexical items. They are thus used to the appearance and use of structures in English. Conversely, they are continually meeting new lexical items or old items used in new ways, and are consequently less likely to identify a deviant lexical item as deviant, judging it instead a usage they have not met before.

The results for L1 language teachers are rather less susceptible of explanation. I suspect, however, that it results from a combination of (a) the concentration on structure in their teaching methods, to the detriment of lexis, and, (b) their previous exposure to L2 usage which may have made them more tolerant of lexical deviance.

Appendix: the sentences used

- A. After the doctor saw him, instead of him feeling better, he feels worse.
- B. After the doctor saw him, instead of him feeling the better, he felt worse.
- C. After the doctor saw him, instead of him feeling better, he felt worse.
- D. After the doctor saw him, instead he felt better, he felt worse.
- E. After the doctor saw him, instead of him feeling lighter he felt worse.

BIBLIOGRAPHY

- Maxwell, A.E. 1961. Analysing qualitative data. Methuen.
- Stuart, A. (1951) "An application of the distribution of the ranking concordance coefficient", Biometrika, 38, pp. 33-42.
- Weatherhead, J.R. (1973) An attempt to ascertain native - speaker ranking of types of overt deviance within English sentences . Unpublished MA thesis, University of Lancaster.