

A51

Jones Glyn

Lexical cohesion as a predictor of item difficulty

It is notoriously difficult to predict the facility value of receptive skills (comprehension) items. Experienced teachers and testers can generally assign a difficulty level to a text with some confidence, but similar judgments about items – the difficulty level of multiple choice question Y *about* text X – are much less reliable.

Macmillan (2007) proposes a framework for classifying multiple choice reading comprehension items by analyzing the link between the task and the source text in terms of lexical cohesion. She posits a typology of links ranging from simple repetition (where the same word forms occur in text and task) to complex relations involving synonymy, hyponymy and ellipsis.

The aim of the present research is to test the hypothesis that item difficulty is related to the number and type of such links. The reading comprehension items from the Pearson Test of English (PTE) all have a known difficulty value established empirically by IRT measurement following field testing by over 200 trial subjects. A random set of these items will be analyzed in accordance with Macmillan's scheme. Factor analysis will be applied in order to identify the aspects of cohesion that correlate most strongly with item difficulty.

It is hoped that the results of the study will form the basis of a method for predicting the difficulty level of comprehension tasks. This in turn could be used to pre-assign new items to trial tests, and to train item writers to rough-tune new items to a particular level.