

Eric Grinstead
 Centralinstitut for nordisk asienforskning
 København

A METHOD FOR THE STUDY OF COMPOUND WORDS IN CHINESE

Lexicostatistical work usually treats script letters as the basic unit. This is not suitable for Chinese, where the basic unit contains a meaning, a pronunciation, and a form.

The "word" we study is actually of two signs, so we must study two-element words. I have tried to find properties that are independent of form, meaning, and pronunciation, but incline to consider the meaning important when analysis is possible.

In a given vocabulary, an element can occur first or second. The number of occurrences before and the number after form a pair that can be represented as a Cartesian point, or as prefer it, a vector.

A two-element compound is defined by the line joining POST to HUS, in vector terms POST minus HUS.

I am only beginning to see what the consequences are, but so far a sample of the less common characters gives a good correlation with the rare characters of a new dictionary published in China, even though the criteria for selection are based only on compounds in my list, and on occurrences in Classical Chinese in the dictionary.

