GLOSSARY

Basic literacy level. In the Lao National Literacy Survey (Lao People's Democratic Republic 2004) an individual was said to be at the Basic literacy level if they obtained 8 or more marks in each of three literacy tests, namely, reading, writing and numeracy.

Boxplot. Also known as 'Box and Whisker Plot', this is a method of graphically displaying the 'Five Figure Summary' of a set of data, namely Maximum, Minimum, Median, Upper Quartile and Lower Quartile.

Cognate. This refers to words or speech forms that share a common ancestral root.

Confidence interval. This is an estimated range of values with a known and controllable probability (typically 0.95) of containing the true value of the population parameter of interest (e.g. literacy rate).

Contingency table. This is a table of counts which describe the relationship between two or more variables. The variables are typically categorical, such as Gender or Nationality. Contingency tables are also known as 'cross-tabulations' or 'cross-classifications'.

Dead syllable. This is a syllable ending in one of the stops *p*, *t*, *k,* or *ʔ*. Gedney (1972:424) used the term 'checked' as an alternative to 'dead'.

Dendogram. This is any type of branching diagram.

Dotplot. This is a method of graphically displaying a set of data in which each individual is represented by a dot. For large data sets, each dot represents a fixed number of individuals.

Fisher Exact test. This test calculates the exact probability of a given 2×2 table of cross-classified frequencies (or any more extreme distribution of cell frequencies) under the assumption that no relationship exists between the row and column variables.

Freeman-Halton test. This is an extension of the Fisher Exact test to cross-classified frequency tables of dimension greater than 2×2.

Functional literacy level. In the Lao National Literacy Survey (Lao People's Democratic Republic 2004) an individual was said to be at the Functional literacy level if they obtained 14 or more marks in each of three literacy tests, namely, reading, writing and numeracy.

Histogram. This is a method of graphically displaying a set of data by means of rectangles whose widths represent class intervals and whose areas are proportional to the corresponding frequencies.

IPA. The International Phonetic Alphabet (IPA) will be used to transcribe language data. Phonetic transcriptions will be marked by square brackets e.g. [fənɛtɪkʰ]; phonemic transcriptions will be marked by forward slashes e.g. /fonimɪkʰ/.

Linear Correlation Coefficient. This is a measure of the strength and direction of linear relationship between two variables. The frequency distributions of the variables are assumed to be from a Normal distribution.

Live syllable. This is a syllable ending in a voiced sound, i.e. a vowel, semi-vowel or nasal. Gedney (1972:424) used the term 'smooth' as an alternative to 'live'.

Log-linear analysis. This is a statistical technique for analysing the data in a contingency table. The effect of several independent variables on a dependent variable is estimated.

Open syllable. This is a syllable ending in a vowel.

Phenogram. This is a branching diagram that links entities by estimates of overall similarity. This is often constructed by using UPGMA (Unweighted Pair-Group Method using Arithmetic Average).

Spearman Rank Correlation. This is a technique for measuring the strength and direction of relationship between two variables. No assumptions are required concerning the frequency distributions of the variables.

Sustained Functional literacy level. In the Lao National Literacy Survey (Lao People's Democratic Republic 2004) an individual was said to be at the Sustained Functional literacy level if they obtained 22 or more marks in each of three literacy tests, namely, reading, writing and numeracy.

t-test. This is a type of statistical hypothesis test in which the test statistic has a Student's t distribution if the null hypothesis is true. There are various types of t-test. A 2-sample t-test is used to test whether or not the mean values of two independent samples are equal. In a paired t-test each member of one sample has a unique relationship with a member of the other sample. The aim is again to test whether or not the mean values of the two samples are equal.

Township. This is an administrative unit in Myanmar immediately above the village tract and below the State or Division. For example, Keng Tung Township is above Yang Lorh Village Tract but below Eastern Shan State.

Unweighted Pair-Group Method using Arithmetic Average (UPGMA). This is a type of cluster analysis.

Village tract or Ward. This is an administrative unit in Myanmar immediately above the village and below the township. For example, Yang Lorh Village Tract is above Yang Lorh Village but below Keng Tung Township.