Multiple and Joint Correspondence Analysis: Testing the True Dimension of a Study

Sergio Camiz

Dipartimento di Matematica, Sapienza Università di Roma

E-mail: sergio.camiz@uniroma1.it,

Gastão Coelho Gomes

Departamento de Métodos Estatísticos,

Universidade Federal do Rio de Janeiro

E-mail: gastao@im.ufrj.br

April 23, 2013

Resumo:

Neste trabalho o problema de propor uma dimensão para Análise de Correspondências Multiplas (MCA) foi discutido em duas direções: a re-avaliação baseada na inércia explicada no sentido de Benzécri (1979) and Greenacre (2006) e um teste proposto por Ben Ammou and Saporta (1998). Isto se faz considerando uma melhor reconstrução dos elementos fora da diagonal da sub-tabela de Burt cruzando as variáveis Nominais. Então Greenacre (1988) introduziu a "Joint Correspondence Analysis" (JCA). Resultados de duas aplicações são apresentados para avaliar a qualidade da reconstrução de ambas MCA e JCA, também são comparados com os resultados de Análise de Correspondências Simpes de tabelas 2 por 2. Observamos que a redução

de dimensão é muito melhor para a técnica JCA do que para a MCA, que se revela viesada e não monótona.

Este trabalho feito durante a visita do prof. Camiz a UFRJ, em abril 2013, apoiada pela FAPERJ (processo APV-E-26/110.018/2013), foi enviado para ser publicado na revista francesa "Revue des Nouvelles Technologies de l'Information (RNTI)".

Abstract

The problem of the proper dimension of a Multiple Correspondence Analysis (*MCA*) is discussed, based on both the re-evaluation of the explained inertia sensu Benzécri (1979) and Greenacre (2006) and a test proposed by Ben Ammou and Saporta (1998). This leads to the consideration of a better reconstruction of the off-diagonal sub-tables of the Burt's table crossing the nominal characters taken into the account. Thus, Greenacre (1988) Joint Correspondence Analysis (*JCA*) is introduced and the results obtained on two applications are shown and the quality of reconstruction of both *MCA* and *JCA* solutions are compared to the Simple Correspondence Analysis results of the two-way tables. It results that *JCA*'s reduced-dimensional reconstruction is much better than the *MCA*'s one, that reveals highly biased and non-monotonous.
Keywords: Correspondence Analysis, Multiple Correspondence Analysis, Joint Correspondence Analysis.

1 Introduction

The identification of the dimension of a data table under study is a crucial issue of most multidimensional scaling techniques. A distinction should be done between linear scaling, in which the encapsulated solutions allows an *a posteriori* choice of the user, and non-linear one, in which usually the solution dimension is an *a priori* choice that conditions the results. As the latter may be only hypothesized,

e.g. according to the results of a previous linear scaling that may be used as a starting configuration, the identification in the linear case has an importance that goes beyond the simple linear case, to involve most of the analysis that follow the scaling itself. To quote only some, the number of factors to be interpreted, those on which to attempt a classification, the dimension in which search for a non-linear solution or for a factor analysis, etc., are all items that depend on this choice.

In this paper, we deal with this problem in the framework of Multiple Correspondence Analysis (*MCA*, Benzécri et al., 1973-82; Greenacre, 1983; Langrand and Pinzón, 2009) in particular considering its alternative, the Joint Correspondence Analysis (*JCA*, Greenacre, 1988), whose solution depends on an *a priori* selected dimensionality, and the partial reconstruction of the original data that results by the application of both *MCA* and *JCA* reconstruction formulas.

The application of these methods to two examples taken from studies in linguistics (Nardi, 2007; Senna, 2013) will show unexpected results when comparing the reconstruction: even if *JCA* was supposed to perform better, the results of *MCA*, in comparison with those of *JCA*, would seriously get questionable its use. Indeed, the application to the Burt's table of the chi-square metrics, and the following correspondence analysis, emphasize too much the importance of the block-diagonal matrices, whose interest is practically null, in respect to the off-diagonal ones that contain the most interesting information.

2 Theoretical framework

In exploratory multidimensional scaling the identification of the proper dimension of the solution is strictly tied to the crucial distinction between relevant and non-relevant information, something similar to the identification of errors in classical statistics, but not the same. In this case, the relevant information is also tied to the possibility to interpret the factors, according to the paradigms of the method at hand: it may be either the percentage of explained inertia for the metric scaling or the stress for the non-metric one, these being in practice the most widely used.

Thus, to take into account a large share of inertia or reduce as much as possible the stress are the most evident rough methods that may be used and a higher-dimensional solution is normally preferred to a smaller one only if these values are significantly smaller. But how to evaluate to what extent they are "significantly smaller"? According to the method at hand, a solution may be found: for Principal Component Analysis, Jackson (1993) compared some of the existing ones in literature.

2.1 Singular Value Decomposition and Generalized Singular Value Decomposition

We may ground our further discussion on the well known Singular Value Decomposition (*SVD*, Greenacre, 1983; Abdi, 2007) theorem, that states

Theorem 1. *Any real matrix X may be decomposed as $X = U\Lambda^{1/2}V'$, with Λ the diagonal matrix of the real non-negative eigenvalues of XX', U the orthogonal matrix of the corresponding eigenvectors, and V the matrix of eigenvectors of $X'X$ (with the same eigenvalues), with both constraints $U'U = I$ and $V'V = I$.*

This theorem corresponds to the reconstruction formula of an r-rank matrix

$$x_{ij} = \sum_{\alpha=1}^{r} \sqrt{\lambda_\alpha} \;\; u_{i\alpha} \; v_{j\alpha}$$

on which the Eckart and Young (1936) theorem is based:

Theorem 2. *(Eckart and Young) The s-rank reconstruction of any real matrix X, with $s < r$, the rank of X, once its singular values are sorted in decreasing order,*

$$x_{ij} \approx \sum_{\alpha=1}^{s} \sqrt{\lambda_\alpha} \;\; u_{i\alpha} \; v_{j\alpha}$$

is the best one in the least-squares sense.

Thus, the exploratory analysis paradigm states that the most relevant information is tied to the largest eigenvalues and the non-relevant to the least ones. The problem of distinguishing among them, that is to identify at least a tentative cutpoint of either the singular- or the eigen-values sequence, remains a crucial issue, that seems more easily solved in the case of Simple Correspondence Analysis (*SCA*, Benzécri et al., 1973-82; Greenacre, 1983; Langrand and Pinzón, 2009), since the special chi-square metrics adopted allows some useful solutions and an easy interpretation of the results.

Indeed, for our purposes, we shall refer to the Generalized Singular Value Decomposition (*GSVD*, Greenacre, 1983; Abdi, 2007). For a given matrix X, this involves using two positive definite square matrices expressing constraints imposed respectively on the rows and the columns of X. If M and N are such matrices, the $GSVD$ aims at decomposing X as $X = U\Lambda^{1/2}V'$, under the orthogonality constraints $U'MU = I$ and $V'NV = I$. We shall express these conditions by saying that U and V are required to be M- and N-orthogonal, respectively.

Theorem 3. *Given two real positive definite matrices M and N, any real matrix X may be decomposed as $X = \widetilde{U}\Lambda^{1/2}\widetilde{V}'$, under constraints $\widetilde{U}'M\widetilde{U} = I$ and $\widetilde{V}'N\widetilde{V} = I$.*

The solution is given by the SVD of the matrix $\widetilde{X} = M^{1/2}XN^{1/2} = F\Lambda^{1/2}G'$, with $F'F = I$, $G'G = I$, $\widetilde{U} = M^{-1/2}F$, and $\widetilde{V} = N^{-1/2}G$. It results that $\widetilde{U}\widetilde{U}' = M^{-1}$ and $\widetilde{V}\widetilde{V}' = N^{-1}$ respectively.

2.2 Correspondence Analysis

Let N an $r \times c$ contingency table, with $n = n_{..}$ the table grand total, $\vec{r} = (p_{1.}, ..., p_{r.})'$ the vector of row marginal profile (with $p_{ij} = n_{ij}/n$), $\vec{c} = (p_{.1}, ..., p_{.c})'$ the vector of column marginal profile, and $D_r = \mathrm{diag}(\vec{r}), D_c = \mathrm{diag}(\vec{c})$ the corresponding diagonal matrices. The *SCA* of N results from the application of *GSVD* to the contingency table N with the constraints given by the diagonal matrices D_r and D_c. As a result,

the reconstruction formula of N is:

$$n_{ij} = nr_i c_j \left(1 + \sum_{\alpha=1}^{\min(r,c)-1} \sqrt{\lambda_\alpha}\ f_{i\alpha}\ g_{j\alpha} \right).$$

This results from the formulation of the problem in terms of the best weighed least-squares approximation of the matrix N by another matrix H of lower rank which minimizes

$$\sum_{i=1}^{r}\sum_{j=1}^{c} \frac{(n_{ij}-h_{ij})^2}{e_{ij}} = \sum_{i=1}^{r}\sum_{j=1}^{c} \frac{(n_{ij}-h_{ij})^2}{nr_i c_j} = n^{-1}\mathbf{trace}\left(D_r^{-1}(N-H)D_c^{-1}(N-H)'\right) \qquad (1)$$

where the weights are the inverse of the expected frequencies. Thus, the reconstruction formula may be well synthesized as

$$N = n\ \vec{r}\ \vec{c}\,' + D_r F \Lambda^{1/2} G' D_c. \qquad (2)$$

As a matter of fact, in order to produce a simultaneous graphical representation, *SCA* eigenvectors are usually rescaled, by defining as *coordinates* the quantities $\Phi = F\Lambda^{1/2}$ and $\Psi = G\Lambda^{1/2}$. With this transformation, and applying the Eckart and Young's theorem, any reduced rank approximation obtained by limiting the sum above to the r largest eigenvalues is the best approximation in the weighed least-squares sense:

$$n_{ij} \approx nr_i c_j \left(1 + \sum_{\alpha=1}^{r} \frac{1}{\sqrt{\lambda_\alpha}}\ \phi_{i\alpha}\ \psi_{j\alpha} \right).$$

It results that the inertia along each dimension α equals ${\chi_\alpha}^2 = n\lambda_\alpha$. As in *SCA* the eigenvalues sum, up to the grand total, to the table chi-square, namely

$$\chi^2 = n \sum_{\alpha=1}^{\min(r,c)-1} \lambda_\alpha,$$

the cutting problem is simply solved by using the classical test for goodness of fit (Kendall and Stuart, 1961) or more easily through the Malinvaud (1987) test. The test may be applied, as, for each α-dimensional partial reconstruction, the residuals

correspond to

$$Q_\alpha = \sum_{ij} \frac{(n_{ij} - \widetilde{n}_{\alpha ij})^2}{\widetilde{n}_{\alpha ij}},$$

asymptotically chi-square-distributed with $(r - \alpha - 1) \times (c - \alpha - 1)$ degrees of freedom. In the formula, $\widetilde{n}_{\alpha ij}$ is the cell value estimated by the α-dimensional solution, and the table chi-square test results when α= 0 and $\widetilde{n}_{0ij} = \frac{n_{i\cdot}\ n_{\cdot j}}{n_{\cdot\cdot}}$ is the expected value under independence. Now, Malinvaud (1987) showed that, by substituting the estimated cell values with the expected ones under independence hypothesis, the formula may be approximated by

$$\widetilde{Q}_\alpha = \sum_{ij} \frac{(n_{ij} - \widetilde{n}_{\alpha ij})^2}{n r_i c_j} = \chi^2 - \sum_{\beta=1}^{\alpha} \chi^2_\beta = n \sum_{\gamma=\alpha+1}^{\min(r,c)-1} \lambda_\gamma,$$

that may be more easily used to check for nullity of the residuals. It is interesting to observe that to the same property may be associated the partial chi-square test for significance associated to each eigenvalue, $\chi^2_\alpha = n_{\cdot\cdot}\lambda_\alpha$, with $df = (r + c - 2\alpha - 1)$ (Kendall and Stuart, 1961), to detect if there are linear ordinations of both rows and column levels that explain the deviation from expectation (Orlóci, 1978). Whereas Malinvaud's is an overall test, that may be used to reject the hypothesis of the residuals randomness, thus suggests to go further in the factors inspection, this test informs on the existence of a significant one-dimensional relation among the rows and column levels, independent from the previous ones. Indeed, non-linear relations may results from the co-occurrence of several one-dimensional solutions (not necessarily significant), as could be the case of the application in section 3.

2.3 Multiple Correspondence Analysis

It is well known that ***MCA*** is but a generalization of ***SCA*** and it is based on ***SCA*** of either the indicator matrix Z, whose rows are the units and the columns are all the levels of the considered variables, or the so-called Burt's table $B = Z'Z$ that gathers all contingency tables obtained by crosstabulating all the variables in Z, including the diagonal tables obtained by crossing each variable with itself. We drop here

other definitions and formulas of both ***SCA*** and ***MCA*** and their relations, that may be found, e.g., in Greenacre (1983) or in Langrand and Pinzón (2009). Suffice here to remind that, in both cases, the chi-square metrics is adopted so that the interpretation of results ought to be done once again in terms of deviations from expectation. It is easy to see that in this case the total inertia of Z is $I_z = \frac{J-Q}{Q}$, where Q is the number of variables and J the total number of levels, that is $J = \sum_{i=1}^{Q} l_i$ where $l_i$is the number of levels of the i-th character and that the eigenvectors in ***SCA*** of both Z and B are the same, whereas the B's eigenvalues are the squares of Z's: $\mu_\alpha^2 = \nu_\alpha$. Thus, it makes no difference to perform ***MCA*** on either matrix.

As ***SCA***, given a Burt matrix B, ***MCA*** may be defined as the weighted least-squares approximation of B by another matrix H of lower rank, minimizing

$$n^{-1}Q^{-2}\mathbf{trace}\left(D_r^{-1}(B-H)D_r^{-1}(B-H)'\right). \tag{3}$$

Notice how (3) derives from (1). In terms of the subtables, this may be rewritten as

$$\begin{aligned} & n^{-1}\mathbf{trace}\left(D^{-1}(B-H)D^{-1}(B-H)'\right) = \\ = & \; n^{-1}\sum_{i=1}^{Q}\sum_{j=1}^{Q}\mathbf{trace}\left(D_i^{-1}(N_{ij}-H_{ij})D_j^{-1}(N_{ij}-H_{ij})'\right), \end{aligned}$$

where H is the supermatrix of the H_{ij}. Introducing the norm notation

$$\|A-B\|_{ij}^2 = \mathbf{trace}\left(D_i^{-1}(A-B)\; D_j^{-1}\; (A-B)'\right)$$

the minimization can be written as

$$n^{-1}\sum_{i=1}^{Q}\sum_{j=1}^{Q}\|N_{ij}-H_{ij}\|_{ij}^2. \tag{4}$$

In ***MCA*** the identification of the true dimension is particularly difficult, despite the ***MCA*** is a ***SCA*** of a particular table, because the chi-square test has no sense. Indeed, for B a chi-squared statistic may again be calculated as if it were a

contingency table, and this simplifies as

$$\chi^2_B = 2\sum_{i=1}^{Q}\sum_{j=1}^{i-1}\chi^2_{ij} + n(J-Q),$$

where χ^2_{ij} is the chi-squared statistic for the off-diagonal subtable $N_{ij} = Z_i'Z_j$ crossing the i-th and the j-th characters, but without the possibility to make a test. Unfortunately neither Q_α nor $\widetilde{Q}_\alpha$ computed on the indicator matrix Z are chi-square distributed (Ben Ammou and Saporta, 1998), since Z is composed by 0's and 1's.

Thus, the only useful information appears to be the tie of *MCA* with Generalized Canonical Analysis (*sensu* Carroll, 1968; Carrol *et al.*, 1986). Indeed, when *MCA* is seen as the analysis of a multi-indicator matrix, the square roots of the eigenvalues may be seen as the sum of the squares of correlations of the corresponding eigenvector with its projections onto the subspaces spanned by the levels of each character. Albeit it may be interpreted as a degree of coherence in the meaning of each projection, this property is very difficult to handle, so that its use is very limited. In practice, the current users are satisfied when the first two-tree factors are enough larger than the following, regardless of their numerical value or of the percentage of cumulated inertia, that is generally admitted to be highly underestimated.

The term "inflation" applied to the high number of eigenvalues of the *MCA*, derives from Benzécri (1979) that explains it in terms of the arbitrary number of levels in wich a continuous character may be discretized to become qualitative and the fact that, if we compare *SCA* and *MCA* applied to the same two characters contingency table, a relation between the eigenvalues may be found. Indeed, by partitioning a two-characters Burt's table $Z'Z$ into submatrices it can be shown (ibid.) the relation $\mu_\alpha = \frac{1\pm\sqrt{\lambda_\alpha}}{2}$ that holds among the eigenvalues of Z and those of the *SCA* of the contingency table crossing the two characters. In this case, it is evident that to the eigenvalues $\lambda_\alpha = 0$ of *SCA* correspond eigenvalues $\mu_\alpha = \frac{1}{2}$ of Z and $\nu_\alpha = \frac{1}{4}$ of B, whereas to the others two correspond, one of which larger and the other smaller than $\frac{1}{2}$ and $\frac{1}{4}$ respectively. Generalizing this argument to several

characters results in admitting to limit attention in *MCA* only to the eigenvalues larger than their mean, that is $\mu \geq \overline{\mu}_\alpha = \frac{1}{Q}$.

The argument is discussed in detail by both Benzécri (1979) and Greenacre (1988, 2006). Both authors suggest, in order to get a measure of relative importance of each factor, to re-evaluate the eigenvalues larger than the mean (equal to $\frac{1}{Q}$) according to the formula

$$\rho\left(\mu_\alpha\right) = \left(\frac{Q}{Q-1}\right)^2 \left(\mu_\alpha - \overline{\mu}\right)^2, \quad \mu_\alpha \geq \overline{\mu} = \frac{1}{Q}.$$

Thus, as Benzécri bases his argument on the discretization of a continuous character, he suggests to consider as total inertia the sum of the re-evaluated eigenvalues and consider as percentage of explained inertia the ratio $\frac{\rho(\mu_\alpha)}{\sum_\alpha \rho(\mu_\alpha)}$. This results in a dramatic re-evaluation of the relative importance of the first eigenvalues. On the opposite, Greenacre bases his arguments on the unusefulness to take into account the diagonal block matrices and the utility to limit attention only to the total off-diagonal inertia of the table, that is the sum of squared (non-re-evaluated) eigenvalues minus the diagonal inertia: that is

$$\frac{Q}{Q-1}\left(\sum_{\mu_\alpha > 1/Q} \mu_\alpha^2 - \frac{J-Q}{Q^2}\right).$$

Experiments show that the Greenacre's reevaluation is always limited to a share of the total inertia of Burt's table even by taking into account all the eigenvalues larger than the mean.

An alternative is proposed by Ben Ammou and Saporta (1998, 2003): they suggest to estimate the significance of the eigenvalues of *MCA* according to their distribution. If the characters are independent, $\sum_{\beta=1}^{J-Q} \mu_\beta = \frac{J-Q}{Q}$ and $S_{\mu^2} = \sum_{\beta=1}^{J-Q} \mu_\beta^2 = \frac{J-Q}{Q^2} + \frac{\sum_{i \neq j} \phi_{ij}^2}{Q^2}$ with $n_{..}\phi_{ij}^2 \approx \chi^2_{(l_i-1)(l_j-1)}$, thus,

$$E[n_{..}\phi_{ij}^2] = E[\chi_{ij}^2] = (l_i - 1)(l_j - 1)$$

so the expectation of the variance S_μ^2 of the eigenvalues is

$$\sigma^2 = E[S_\mu^2] = \frac{1}{n_{..}Q^2(J-Q)} \sum_{i \neq j} (l_i - 1)(l_j - 1).$$

Roughly, one may assume that the interval $\frac{1}{Q} \pm 2\sigma$ should contain about 95% of the eigenvalues. Indeed, since the kurtosis of the set of eigenvalues is lower than for a normal distribution, the actual proportion is larger than 95%.

2.4 Joint Correspondence Analysis

Greenacre (1988) criticizes *MCA* approach since in his opinion "it is not a natural generalization of the geometrical [...] or the least squares approach [of *SCA*]" and proposes his *Joint Correspondence Analysis* (*JCA*) as its natural generalization to the case of nominal data, considered as a set of contingency tables obtained by crossing them on the same individuals. According to him, in *MCA* "appears to be no justification for fitting the diagonal subtables B which contribute the term $n(J-Q)$ to the total variation", a term that "artificially inflates the total variation to the extent that the percentages accounted for by the major principal axes can be very low, especially when $J-Q$ is large. A more natural measure of total variation is the sum $\sum\sum_{q \neq s} \chi_{qs}^2$. This suggests an alternative generalization of correspondence analysis which fits only the off-diagonal contingency tables, analogous to factor analysis where values on the diagonal of the covariance or correlation matrix are of no direct interest."

Indeed, the proposed redefinition of the total variation, by removing the diagonal block-matrices, would fix an important bias due to the application to the Burt's table of the chi-square metrics, as the diagonal structure of the diagonal block-matrices represents a very high deviation from the expected values, that *MCA* analyzes as if it were a true deviation. On this basis, on the opposite to the current use, this kind of analysis is not really suitable.

So, Greenacre (1988) proposes his *Joint Correspondence Analysis* (*JCA*) as a

weighed least-squares approximation aiming at minimizing

$$n^{-1}\sum_{i=1}^{Q}\sum_{j=1}^{i-1}\|N_{ij}-H_{ij}\|_{ij}^{2}\,, \tag{5}$$

instead of (4) **with the corresponding** $\chi_J^2=\sum_{i=1}^{Q}\sum_{j=1}^{i-1}\chi_{ij}^2$, **sum of the chi-squares of all off-diagonal tables, that unfortunately may not be checked for significance.**

In order to get the solution, he proposes an alternating least-squares algorithm, based on the reformulation of (5) **as follows:**

$$n^{-1}\sum_{i=1}^{Q}\sum_{j=1}^{i-1}\|N_{ij}-H_{ij}\|_{ij}^{2}=n^{-1}\sum_{i=1}^{Q}\sum_{j=1}^{i-1}\left\|N_{ij}-n\;\vec{r}_i\;\vec{r}_j\;'-L_{ij}\right\|_{ij}^{2} \tag{6}$$

with $\vec{r}_i$ **the diagonal of the** i**-th block-diagonal matrix. Calling** H **and** L **the supermatrices gathering the** H_{ij} **and** L_{ij} **respectively, Greenacre (1988) states the equivalence of the rank-**K **solution of** L **which satisfies the normal equations in the minimization of the second term of** (6) **with the rank-**$(K+1)$ **matrix** $H=\vec{r}\;\vec{r}\;'+L$ **which satisfies minimizing** (5)**, with** $\vec{r}$ **the supervector gathering the** Q **vectors** $\vec{r}_i$**.**

The matrix approximation L **of rank** K **is of the form** $L=nDXD_\beta X'D$**, where the** $J\times K$ **matrix** X **is normalized as** $X'DX=QI$**, with** $D=\mathbf{diag}(\vec{r})$**. The matrix** X **of parameters has rows corresponding to the categories of the variables and columns corresponding to the dimensions of the solution, that must be chosen in advance. The diagonal matrix** D_β **contains a scale parameter for each dimension. This form of** L **and the normalization conditions are chosen to generalize the bivariate case** (2)**. The parameter matrix** X **is partitioned row-wise according to the variables as** $X_1,\cdots,X_Q$**, where** X_q **is** $J_q\times K$**, so that the submatrices of** L **are** $L_{qs}=nD_qX_qD_\beta X_s'D_s$**. There are also inherent centering constraints on** X **of the form** $X'r=0$ **due to the orthogonality with the dimension defined by the trivial solution. It is evident that the dimension of the solution must be chosen in advance.**

It is to be noted that fitting the off-diagonal submatrices reminds the *MINRES* method for least-squares factor analysis where the off-diagonal elements of a correlation matrix are fitted (Thomson, 1934, see also Gabriel, 1978).

Thus Greenacre (1988) proposes the approximate reconstruction of the whole matrix $B - n\ \vec{r}\ \vec{r}\ '$, namely

$$B - n\ \vec{r}\ \vec{r}\ ' \approx nDXD_{\beta}X'D + C,$$

where C is a block diagonal matrix with submatrices $C_{qq}, \quad q = 1, ..., Q$ down the diagonal and zeros elsewhere. Here, each C_{qq} is composed by dummy parameters which effectively allow perfect fitting of the submatrices on the diagonal of $B - n\ \vec{r}\ \vec{r}\ '$, thereby eliminating their influence on the model of interest. The minimization of

$$\begin{aligned} B - n\ \vec{r}\ \vec{r}\ ' = 2n^{-1} \sum_{i=1}^{Q} \sum_{j=1}^{i-1} \left\| N_{ij} - n\ \vec{r}_i\ \vec{r}_j\ ' - L_{ij} \right\|_{ij}^2 \\ +\ \ n^{-1} \sum_{k=1}^{Q} \left\| N_{kk} - n\ \vec{r}_k\ \vec{r}_k\ ' - L_{kk} - C_{kk} \right\|_{k}^2 . \end{aligned} \tag{7}$$

is equivalent to minimizing (6) because the latter set of terms in (7) can always be made zero by setting $C_{ii} = N_{ii} - n\ \vec{r}_i\ \vec{r}_i\ ' - L_{ii}$.

The algorithm proposed by Greenacre (1988) to minimize (7) can be performed iteratively by alternating between the variables in C and those in X and D_{β} as follows:

1. **fix the dimension K of the solution.**

2. **initiate the algorithm with an analysis of the full Burt matrix B, that is**

$$B - n\ \vec{r}\ \vec{r}\ ' \approx nDXD_{\beta}X'D. \tag{8}$$

3. **limiting attention to the first K dimensions, say the first K columns of X $\vec{x}_{(1)}, \cdots, \vec{x}_{(K)}$, (8) can be rewritten as**

$$B - n\ \vec{r}\ \vec{r}\ ' \approx \sum_{k=1}^{K} n\beta_k D\vec{x}_{(k)}\vec{x}\ '_{(k)}D.$$

so that, if all quantities except the β_k ($k = 1, \cdots, K$) are regarded as fixed, the

problem reduces to a simple weighted least-squares regression (see Greenacre, 1988, for further details).

4. **Keeping X and D_β fixed, set**

$$C_{ii} = N_{ii} - n\ \vec{r}_i\ \vec{r}_i\ ' - nD_iX_iD_\beta X_i'D_i \qquad (i = 1, \cdots, Q).$$

5. **Keeping C fixed, minimize with respect to X and D_β: this is achieved by performing a correspondence analysis on the table $B* = B - C$, that is the Burt matrix with modified submatrices on its diagonal, setting X equal to the first K vectors of optimal row or column parameters and the diagonal of D_β equal to the square roots of the first K principal inertias respectively.**

6. **Iterate the last two steps until convergence.**

In the special case $\mathbf{Q = 2}$, where the problem reduces to fitting the single off-diagonal submatrix $\mathbf{N}_{12}$, the initial solution described above is optimal and provides the simple correspondence analysis of $\mathbf{N = N}_{12}$ exactly.

3 Two applications

To deal with both examples, all computations have been performed with the *ca* package (Nenadic and Greenacre, 2006, 2007) contained in the *R* environment (R-project, 2009).

3.1 A small example

To show in detail the different behavior of the different correspondence analyses, we refer to a data set taken from Nardi (2007), consisting in 2000 words taken from four different kind of periodic reviews (*Childish (TC), Review (TR), Divulgation (TD),* and *Scientific Summary (TS)*), classified according to their grammatical kind (*Verb (WV), Noun (WN),* and *Adjective (WA)*) and the number of internal

layers (*Two- (L2), Three- (L3),* and *Four and more layers (L4)*), as a measure of the word complexity.

In Table 1 the Burt's table that results by crossing the three characters is reported. In Table 2 are represented the first results of the *SCA*s of the three contingency data tables, crossing the three characters two by two, limited to the first two eigenvalues, namely, the eigenvalues, the percentage of corresponding inertia, and the p-value associated to the chi-square calculated for the corresponding one-dimensional reconstruction, that in this case is identical to the Malinvaud's test, since each solution is 2-dimensional. In two cases, the chi-squares test that the second factor has no real meaning, since the p-value is larger than 5%, whereas for the case of the table crossing the type of publication and the kind of words the second factor is also significant. In Figure 1 the results of the three *SCA*s are represented too: it must be pointed out that the vertical position of the items is significant only for the second graphic. Indeed, the inspection of this factor plane shows an arch pattern due to a Guttman effect (Guttman, 1941; Camiz, 2005).

Running *MCA*, the pattern of eigenvalues is represented in Table 3, in which are reported the singular values of the indicator matrix Z, their percentage to their total (that equals $\frac{J-Q}{Q} = 2.33$), the cumulate percentage, the eigenvalues of the Burt's matrix, corresponding to the inertia explained by the factor, and the cumulate inertia.

Indeed, according to both Benzécri (1979) and Greenacre (1988), only three singular values are larger than $1/Q = 1/3$, so that the re-evaluations, reported in Table 4, are referenced to only three dimensions, albeit the fourth is very close to this value (0.33). In both cases, the first dimension re-evaluated inertia is by far larger than the others.

If we apply the Ben Ammou and Saporta (1998, 2003) estimation of the average singular value distribution under independence, we find that the standard deviation is $\sigma = 0.0159364$, so that the confidence interval at 95% level is $(0.30146 < \lambda < 0.36521)$. As a consequence, only the first singular value is outside the confidence interval and should be considered significant. As a matter of facts, the second one is very

close to the threshold (0.3640): this is consistent with the fact that one of the 2-dimensional tables has a significant second eigenvalue.

Let us look now at the one-dimensional reconstruction, as resulting by the *SCA*s of the three individual tables, by the *MCA*, and by Greenacre's *JCA* as reported in Table 5. The comparison of the *SCA* one-dimensional solutions with the original tables shows that the amount of the cumulate absolute residuals is in good agreement with the quality of the solution, as represented by the corresponding chi-square. For this reason, the low quality of the reconstruction of the table crossing kind of words with the type of publications depends on the significance of the second dimension of the *SCA* of this table. At first glance, it is evident the high difference in the cumulate absolute residuals of *MCA* in respect to the other solutions, that is an important sign of the limits of *MCA* in respect to *JCA*. Indeed, the quality of *JCA* one-dimensional reconstruction is in all cases acceptable, so that it is possible to observe a synthetical graphical representation of the three tables that is realistic. On the opposite, the *MCA* reconstruction is dramatically bad: in Table 6 are reported the cumulate absolute residuals of reconstructions of both *MCA* and *JCA*, both for the whole Burt's table and for the three off-diagonal two-way tables. The residuals for 0-dimension are the deviations from independence and the following are reported for all the allowed dimensions: $7 = J - Q$ for *MCA* and 3 for *JCA*, that corresponds to the number of singular values of the Burt's table larger than the mean. Looking at the table, we may notice a continuous decrease of the total residuals in both analyses, with a perfect fit for the total reconstruction of *MCA*, decrease that is somehow slower for *JCA*. On the opposite, the off-diagonal reconstruction of *JCA* is fast and effective, with the 3-dimensional solution nearly perfect, whereas the reconstruction of *MCA* follows a very different pattern. Indeed, the off-diagonal residuals increase progressively, instead of diminishing, until the average eigenvalue, then lower, but improving the reconstruction in respect to the deviation from independence only with the last two dimensions.

To graphically study the results, we can now compare the 2-dimensional graphics obtained by the three *SCA*s, shown in Figure 1, with those obtained by both *MCA*

and *JCA*, shown in Figure 2. The position of the levels of each character are represented on the plane spanned by the first two factors. Considering also that the second dimension is limited in significance, we may note that both *MCA* and *JCA* factor planes represent a good compromise among the three 2-dimensional graphics. The reciprocal positions of the items are not so different among *MCA* and *JCA*: only *WV* and *TS*, are more shifted and their position on *JCA* plane seemsbetter reflect their relation with the other levels.

3.2 A larger example

This second example is taken from a work in progress concerning the definition of an index for the degree of mental disease of patients affected by aphasia (Senna, 2013). For this aim, 46 patients (half of them not affected, taken as control group) were submitted to a test, in which each one had to identify and verbalize 154 images. In this example we consider six scale characters taken by the resulting data table: two of them, *Time Response* (in blue in the graphics) and *Segments Substitution* (orange), result from the test itself; two, *Frequency* (green) and *Primitiveness* (red), are features of the images and their name; and two, *Time of disease* (black) and *Oral Comprehension* (dark red), concern the patients' conditions. The characters' levels are 4, 5, 4 , 4 , 5 , and 3, respectively, summarizing 25 levels. In this case, the Burt's table is composed by 15 off-diagonal tables and is reported in Table 10.

The *MCA* gives 19 non-zero eigenvalues, of which 8 above the average (0.1667) and only 5 above the 95% confidence interval upper bound (0.1778), assumed by Ben Ammou and Saporta (1998, 2003) as a threshold for the number of factors. In Table 7 the sequence of all the eigenvalues is reported. The inertia re-evaluation is shown in Table 8. Looking at the re-evaluated values, it results that the factors following the third do not add more than 1% of inertia, a too small value to deserve being really taken into account. Note that, according to Benzécri (1979) the three-dimensional representation explains over 98% of total inertia, whereas according

to Greenacre (1988) it is only 74.78% (but indeed 98% of the possible total).

We ran *JCA* on the same table and we can compare the step-by-step reconstruction with *MCA*, as for the other example (see Table 9). Once again are visible both the non-monotonicity of the *MCA* off-diagonal pattern and its tremendous reduction in *JCA*. Concerning the relative importance of the axes within the three-dimensional solution, we may say that the percentage of inertia attributed to them is 60.11, 21.89, and 17.99% respectively. It may be noted that, in respect to the maximum inertia solution obtained, the 8-dimensional one, it represents over 90% of the latter.

Eventually, the pattern of levels of each character on the planes spanned by the factors 1-2 and 1-3 is represented for both *MCA* and for *JCA*. Comparing the two graphics in Figure 3, that is the representation of the trajectories on the factor plane spanned by the axes 1 and 2, it is clearly visible that in *JCA* their relative range is somehow changed. In particular, all of them are enlarged in respect to the *Segment Substitution* one. On this plane, the first factor opposes the lowest levels on the right side (typical of the non-affected control patients) to the highest ones on the left. On the other side, it is difficult to derive an interpretation of the second factor, dominated by the *Segment Substitution* on the upper side side (with the minimum folded) and the *Oral Comprehension* with its intermediate level opposed to both others on the lower side. As well, the *Time of Disease* develops most along this factor, but with a folded pattern. On the following graphics in Figure 4, that represent the pattern on the plane spanned by the axes 1 and 3, the same adjustment results, that indeed gets more interpretable the mutual relations between the characters. On the other side, it is evident the highest agreement of *Time of Response* and *Familiarity* both among themselves and with the third factor, so that they appear really independent from the others. Only for the highest levels of the other characters there is a slight agreement, but folded, thus of difficult interpretation. Similar comments may be done on the planes spanned by the axes 2 and 3 (not shown), that confirm the independence between *Time of Response* and *Familiarity* in respect to all other characters.

4 Conclusion

This study started with the aim to understand to what extent the *JCA* (Greenacre, 1988) could be of help in identifying the true dimension of an analysis concerning a set of qualitative data. In this sense, the confidence interval proposed by Ben Ammou and Saporta (1998, 2003) seems a useful answer to this problem, in agreement with the most one-dimensional solution of the *SCA*s applied to the two-way tables of the first application. During the study, the problem of the data reconstruction not only showed that *MCA* is bad in reconstructing the data table, due to the inflation in the number of eigenelements, but also that the re-evaluations proposed by both Benzécri (1979) and Greenacre (2006) do not take into account the fact that the reconstruction of the two-way off-diagonal tables is for the most reduced-dimensional solutions worst than the initial independence table. This may explain the problem encountered by both Camiz and Ferrazza (2006) and Camiz and Venditti (2007) that needed the whole *MCA* reconstruction to perform a qualitative discriminant analysis *sensu* Saporta (1975) of some quality: indeed, the bad reduced dimensional reconstruction could be the cause of the bad discrimination that resulted by withdrawing the dimensions with lowest inertia. To get closer to the daily use of the graphics, as a help for the description and the interpretation of the data, the higher homogeneity of the ranges of the various characters on factor planes of *JCA* improves the interpretation ability of the graphics themselves. It is very strange that, despite the number of studies developed on *MCA*, no trace results in literature of the serious drawbacks found in *MCA*, nor Greenacre (1988) and the followers (Tateneni and Browne, 2000; Vermunt and Anderson, 2005; Greenacre, 2006) quote their important improvement. Thus, *JCA* seems a most promising development and its properties deserve some further deepening. Acknowledgements

This work was mostly carried out during the reciprocal visits of both authors in the framework of the bilateral agreement between Sapienza Universitá di Roma and Universidade Federal do Rio de Janeiro, of which both authors are the scientific responsible. The first author was also granted by his Faculty of belonging, the

Facoltà d'Architettura ValleGiulia of Sapienza and FAPERJ of Rio de Janeiro. All institutions grants are gratefully acknowledged.

References

Abdi, H. (2007). Singular Value Decomposition (SVD) and Generalized Singular Value Decomposition ($GSVD$). In: N. Salkind (Ed.), *Encyclopedia of Measurement and Statistics*. Thousand Oaks, CA: Sage.

Ben Ammou, S., Saporta G. (1998). Sur la normalité asymptotique des valeurs propres en ACM sous l'hypothèse d'indépendance des variables. *Revue de Statistique Appliquée*, 46(3), 21-35.

Ben Ammou, S., Saporta G. (2003). On the connection between the distribution of eigenvalues in multiple correspondence analysis and log-linear models. *REVSTAT-Statistical Journal*, 1(0), 42-79.

Benzécri, J.P., et coll. (1973-82). *L'Analyse des données*, Tome 2. Paris: Dunod.

Benzécri, J.P. (1979). Sur les calcul des taux d'inertie dans l'analyse d'un questionnaire. *Les Cahiers de l'Analyse des Données*, 4(3), 377-379.

Camiz, S. (2005). The Guttman Effect: its Interpretation and a New Redressing Method. *Tetradia Analushsq Dedomenwn (Data Analysis Bulletin)*, 5, 7-34.

Camiz, S., Ferrazza, E. (2006). Studio sull'iconografia di Aiace Telamonio con metodi di analisi esplorative dei dati. *Archeologia e Calcolatori*, 17, 45-70.

Camiz, S., Venditti, S. (2007). Unsupervised and Supervised Classifications of Egyptian Scarabs Based on the Qualitative Characters of Typology. *Archaeological Computing Newsletter*, 67, 9-17.

Carroll, J.D. (1968). Generalization of canonical correlation analysis to three or more sets of variables. Proceedings of the 76th Annual Convention of the American Psychological Association, 3, 227-228.

Carroll, J.D., Green, P.E., Schaffer, C.M. (1986). Interpoint Distance Comparisons in Correspondence Analysis. *Journal of Marketing Research*, 23(3), 271-280.

Eckart, C., Young, G. (1936). The approximation of one matrix by another of lower rank. *Psychometrika*, 1, 211-218.

Gabriel, K.R. (1978). The complex correlational biplot. In: S. Shye (Ed.), *Theory Construction and Data Analysis in the Behavioral Sciences*. San Francisco: Jossey-Bass, 350-370.

Greenacre, M.J. (1983). *Theory and Application of Correspondence Analysis*. London: Academic Press.

Greenacre, M.J. (1988). Correspondence analysis of mutlivariate categorical data by weighted least squares. *Biometrika*, 75, 457-467.

Greenacre, M.J. (2006). From Simple to Multiple Correspondence Analysis. In: Greenacre and Blasius (2006) (Eds.), 41-76.

Greenacre, M.J., Blasius, J. (Eds.) (2006). *Multiple Correspondence Analysis and Related Methods*. Dordrecht (The Netherlands): Chapman and Hall (Kluwer).

Guttman, L. (1941). The Quantification of a Class of Attributes: a Theory and Method of Scale Construction. In P. Horst (Ed.) *The Prediction of Personal Adjustment*. New York, Social Science Research Council.

Jackson, D.A. (1993). Stopping Rules in Principal Component Analysis: a Comparison of Heuristical and Statistical Approaches. *Ecology*, 74(8), 2204-2214.

Kendall, M.G., Stuart, A. (1961). *The Advanced Theory of Statistics*, vol. 2. London: Griffin.

Langrand, C., Pinzón, L.M. (2009). *Análisis De Datos. Métodos y ejemplos*. Bogotá: Escuela Colombiana de Ingenieria Julio Garavito.

Malinvaud, E. (1987). Data analysis in applied socio-economic statistics with special consideration of correspondence analysis. Marketing Science Conference, Joy en Josas: HEC-ISA.

Nardy, M.N.S. (2007). A sintaxe no interior das palavras - efeitos de gênero na lingua escrita contemporânea. PhD Thesis in Linguistics. Rio de Janeiro, Facultade de Leteras da Universidade Federal de Rio de Janeiro.

Nenadic, O., Greenacre, M. (2006). *Computation of multiple correspondence analysis, with code in R.* In: Greenacre and Blasius (2006) (Eds.), 523-551.

Nenadic, O., Greenacre, M. (2007). Correspondence analysis in R, with two- and three-dimensional graphics: the *ca* package. *Journal of Statistical Software*, 20(3), 1-13.

Orlóci, L. (1978). *Multivariate Analysis in Vegetation Research*, 2nd ed.. Den Haag: Junk.

R-project (2009), http://www.r-project.org/

Saporta, G. (1975). Liaison entre plusieurs ensembles de variables et codage de donnes qualitatives. Thse de troisime cycle, Universit Pierre et Marie Curie, Paris VI.

Senna, F.D. (2013). *Acesso e representaç ão lexical na produção de indivíduos afásicos sob a ótica da fonodiologia de uso.* Ph.D. Thesis in Linguistics. Rio de Janeiro, Facultade de Letras da Universidade Federal de Rio de Janeiro.

Tateneni, K., Browne, M.W. (2000). A noniterative method of joint correspondence analysis. *Psychometrika*, 65, 2, 157-165.

Thomson, G.H. (1934). Hotelling's method modified to give Spearman's *g. J. Educ. Psychol.*, 25, 366-74.

Vermunt, J.K., Anderson, C. (2005). Joint Correspondence Analysis (JCA) by Maximum Likelihood, *European Journal of Research Methods for the Behavioral and Social Sciences*, 1(1), 18-26.

Table 1: *Burt's table of the words' type example.*

	L2	L3	L4	WN	WV	WA	TC	TR	TD	TS
L2	1512	0	0	788	483	241	433	385	399	295
L3	0	375	0	203	23	149	64	82	86	143
L4	0	0	113	62	9	42	3	29	21	60
WN	788	203	62	1053	0	0	229	284	273	267
WV	483	23	9	0	515	0	174	133	125	83
WA	241	149	42	0	0	432	97	79	108	148
TC	433	64	3	229	174	97	500	0	0	0
TR	385	82	29	284	133	79	0	496	0	0
TD	399	86	21	273	125	108	0	0	506	0
TS	295	143	60	267	83	148	0	0	0	498
	L2	L3	L4	WN	WV	WA	TC	TR	TD	TS

Table 2: *SCA of the three contingency data tables of words' type example, crossing the three characters two by two. In the columns, the eigenvalues, the percentage of inertia, and the p-value of the chi-square associated to the factors.*

	words vs. levels			publications vs. words			publications vs. levels		
N.	eigen	%	*p*-value	eigen	%	*p*-value	eigen	%	*p*-value
1	.0925	99.98	.0000	.0253	80.53	.0000	.0619	98.82	.0000
2	.0000	0.02	.8625	.0061	19.47	.0022	.0007	1.18	.4771

Table 3: *MCA singular values, percentage to the total and cumulate percentage, eigenvalues, and cumulate inertia of the Burt's table of words' type example.*

Number	Singular value	Percentage	Cumulate %	Eigenvalue	Cumulate inertia
1	0.4896	20.98	20.98	0.239688	0.239688
2	0.3640	15.60	36.58	0.132472	0.372160
3	0.3434	14.72	51.30	0.117930	0.490090
4	0.3300	14.14	65.44	0.108885	0.598975
5	0.3084	13.22	78.66	0.095100	0.694076
6	0.2728	11.69	90.35	0.074431	0.768507
7	0.2252	9.65	100.00	0.050713	0.819220

Table 4: *Inertia re-evaluation according to both Benzécri (1979) and Greenacre (1988) of words' type example.*

	Benzécri's Re-evaluation			Greenacre's Re-evaluation		
Number	Inertia	%	Cum.%	*Inertia*	%	Cum.%
1	0.0549	95.91	95.91	0.2344	88.36	88.36
2	0.0021	3.69	99.60	0.0460	3.40	91.76
3	0.0002	0.40	100.00	0.0151	0.37	92.13
Total	0.0572	100.00		0.2954	92.13	

Table 5: *Original two-way contingency tables of words' type example and their reconstruction according to the first dimension of* SCA*s,* MCA*, and* JCA*, with the corresponding cumulate absolute residuals.*

Original Burt's Matrix

	WN	WV	WA		TC	TR	TD	TS		TC	TR	TD	TS
L2	788	483	241	L2	433	385	399	295	WN	229	284	273	267
L3	203	23	149	L3	64	82	86	143	WV	174	133	125	83
L4	62	9	42	L4	3	29	21	60	WA	97	79	108	148

SCA First Layer

	WN	WV	WA		TC	TR	TD	TS		TC	TR	TD	TS
L2	788	483	241	L2	435	382	400	296	WN	253	257	267	276
L3	204	23	149	L3	60	89	85	141	WV	165	144	127	79
L4	61	9	42	L4	5	25	22	61	WA	82	96	112	142

SCA cumulate absolute residuals

	2				107					2210			

MCA First Layer

	WN	WV	WA		TC	TR	TD	TS		TC	TR	TD	TS
L2	770	559	183	L2	492	409	401	211	WN	249	257	264	283
L3	216	-24	183	L3	13	69	82	211	WV	219	155	145	-3
L4	67	-20	66	L4	-5	18	23	76	WA	32	84	97	219

MCA cumulate absolute residuals

	14440				18972					21183			

JCA First Layer

	WN	WV	WA		TC	TR	TD	TS		TC	TR	TD	TS
L2	783	484	245	L2	435	391	393	293	WN	259	260	266	269
L3	207	29	139	L3	53	82	87	153	WV	160	136	136	82
L4	63	2	48	L4	12	24	25	52	WA	81	100	104	147

JCA cumulate absolute residuals

	280				488					2570			

Table 6: *Words' type example. Absolute residuals of the reduced dimensional reconstructions of both the Burt's table and the two-way off-diagonal ones according to* MCA *and* JCA *respectively: to 0 correspond the deviations from independence.*

	MCA		JCA	
Dim	total	Off-diag.	total	Off-diag.
0	2052807	50788	2052807	50788
1	1426816	54595	1560452	3338
2	1012894	115712	1451977	1003
3	791539	147734	1379887	21
4	570840	120163		
5	269518	52164		
6	133539	34851		
7	0	0		

Table 7: *Aphasia example: MCA singular values, percentage to the total and cumulate percentage, eigenvalues, and cumulate inertia of the Burt's table.*

Number	Singular value	Percentage	Cumulate %	Eigenvalue	Cumulate inertia
1	0.3831	12.10	12.10	0.146759	0.146759
2	0.2774	8.76	20.86	0.076924	0.223683
3	0.2538	8.01	28.87	0.064394	0.288077
4	0.1951	6.16	35.03	0.038073	0.326150
5	0.1829	5.78	40.81	0.033462	0.359612
6	0.1734	5.48	46.28	0.030081	0.389693
7	0.1729	5.46	51.74	0.029885	0.419578
8	0.1705	5.38	57.13	0.029060	0.448638
9	0.1668	5.27	62.39	0.027825	0.476463
10	0.1655	5.23	67.62	0.027388	0.503851
11	0.1610	5.08	72.70	0.025917	0.529768
12	0.1546	4.88	77.59	0.023903	0.553671
13	0.1467	4.63	82.22	0.021533	0.575204
14	0.1398	4.41	86.64	0.019542	0.594747
15	0.1343	4.24	90.88	0.018035	0.612782
16	0.0928	2.93	93.81	0.008604	0.621386
17	0.0820	2.59	96.40	0.006723	0.628110
18	0.0658	2.08	98.47	0.004328	0.632438
19	0.0484	1.53	100.00	0.002339	0.634777

Table 8: *Aphasia example: inertia re-evaluation according to both Benzécri (1979) and Greenacre (1988)*

	Benzécri's Re-evaluation			Greenacre's Re-evaluation		
Number	Inertia	%	Cum.%	*Inertia*	%	Cum.%
1	0.0674	69.04	69.04	0.2597	52.53	52.53
2	0.0176	18.06	87.09	0.1328	13.74	66.27
3	0.0109	11.18	98.27	0.1045	8.51	74.78
4	0.0012	1.19	99.46	0.0341	0.91	75.69
5	0.0004	0.39	99.85	0.0195	0.30	75.98
6	0.0001	0.07	99.92	0.0081	0.05	76.03
7	0.0001	0.06	99.98	0.0074	0.04	76.08
8	0.0000	0.02	100.00	0.0046	0.02	76.09
9	0.0000	0.00	100.00	0.0002	0.00	76.09
Total	0.0977	100.00		0.5710	76.09	

Table 9: *Aphasia example: absolute residuals of the reduced dimensional reconstructions of both the Burt's table and the two-way off-diagonal ones according to MCA and JCA respectively: to 0 correspond the deviations from independence.*

	MCA		JCA	
Dim	total	Off-diag.	total	Off-diag.
0	84100	17917	84100	17917
1	64651	12369	61508	10275
2	59923	11766	53545	7557
3	48571	7980	41627	3257
4	47619	10017	37899	2255
5	46863	10714	36823	1937
6	46682	11475	34737	1304
7	46134	12377	33401	810
8	44534	13167	32943	685
9	44241	13311	31939	340
10	41003	12687		
11	34973	10431		
12	33437	9963		
13	30953	9617		
14	26018	8555		
15	18406	5441		
16	14641	4559		
17	8992	2963		
18	4357	1341		
19	0	0		

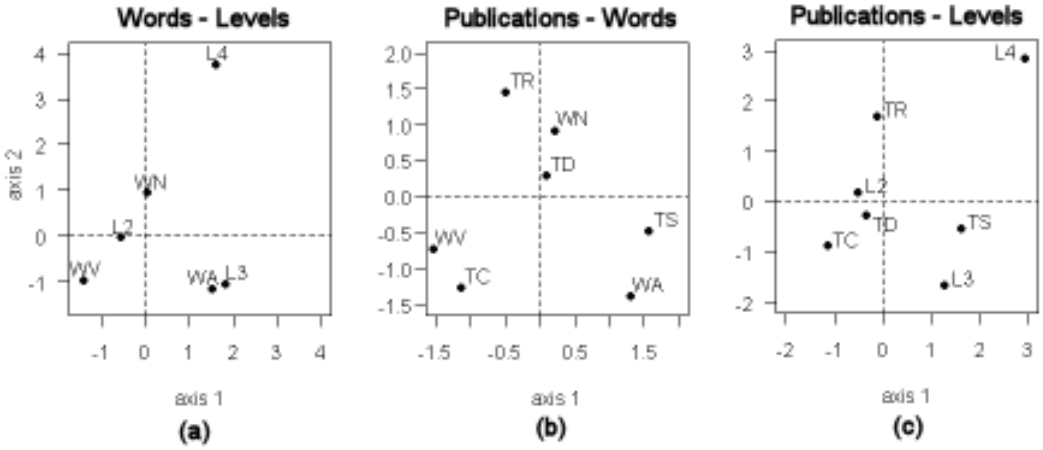

Figure 1: *Words' type example: The pair of characters levels on the three two-way SCAs: (a) Words vs. Levels; (b) Publications vs. Words; (c) Publications vs. Levels.*

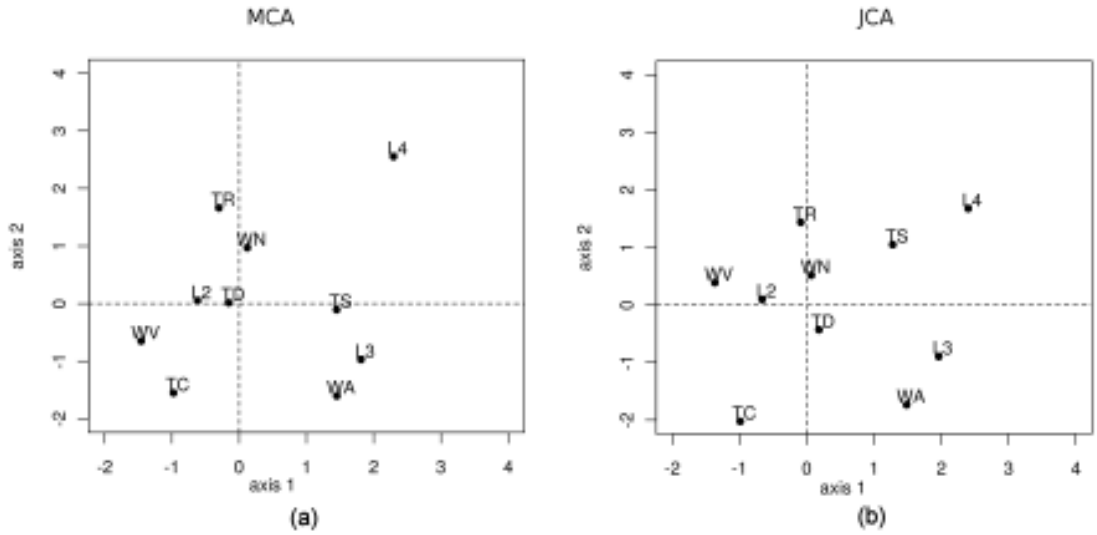

Figure 2: *Words' type example: representation of the three-characters levels on the plane spanned by the first two factors: (a) MCA; (b) JCA.*

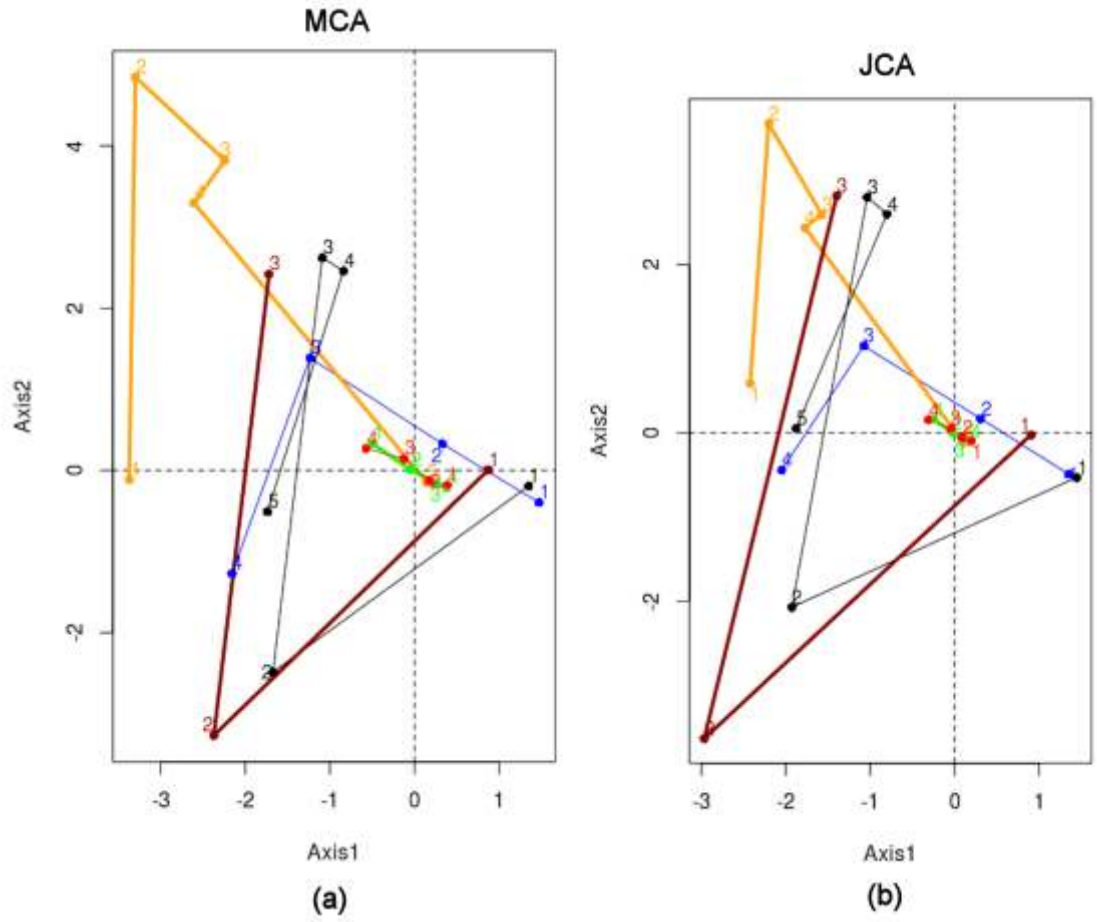

Figure 3: *Aphasia example: representation of the six characters trajectories on the plane spanned by the first two factors: (a) MCA; (b) JCA.*

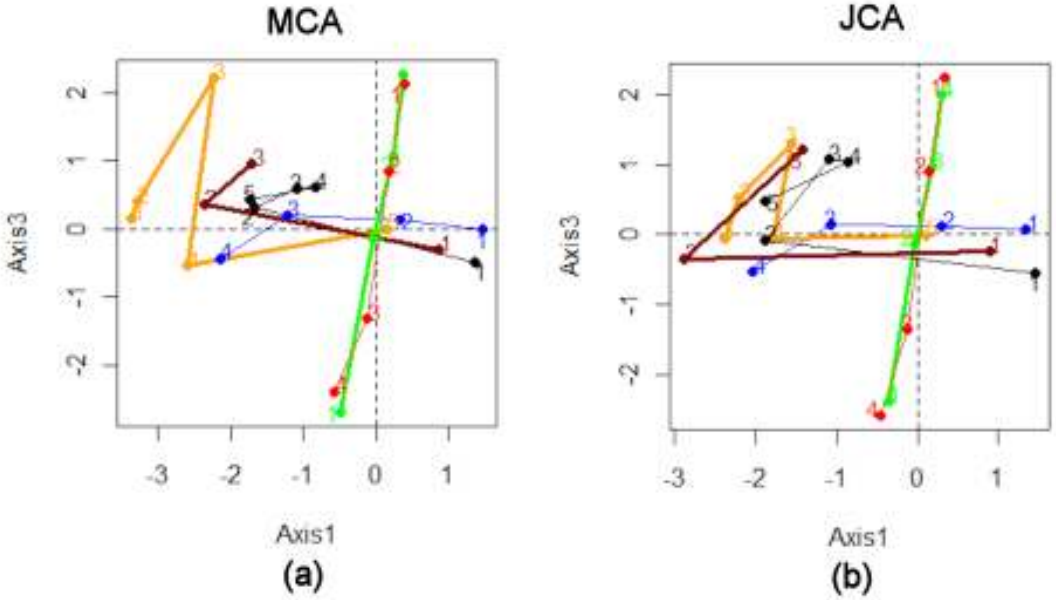

Figure 4: *Aphasia example: representation of the six characters trajectories on the plane spanned by the factors 1 and 3: (a) MCA; (b) JCA.*

Table 10: *Burt's table of the six-characters data set of Aphasia example.*

	Time Response				Segments Substitution					Frequency				Primitiveness				Time of disease					Oral Comprehension		
	1	2	3	4	1	2	3	4	5	1	2	3	4	1	2	3	4	1	2	3	4	5	1	2	3
1	2056	0	0	0	0	0	0	12	2044	336	693	670	357	480	908	403	265	1746	98	21	128	63	1934	36	86
2	0	2756	0	0	8	3	12	73	2660	582	942	799	433	601	1107	495	553	1505	370	115	527	239	2083	205	468
3	0	0	1055	0	26	10	17	100	902	277	400	232	146	202	371	199	283	215	202	100	315	223	547	175	333
4	0	0	0	1217	31	7	1	62	1116	369	449	277	122	143	420	237	417	76	408	72	262	399	364	508	345
1	0	8	26	31	65	0	0	0	0	16	28	14	7	10	20	15	20	1	13	11	14	26	15	26	24
2	0	3	10	7	0	20	0	0	0	6	7	6	1	2	7	3	8	0	4	5	7	4	2	1	17
3	0	12	17	1	0	0	30	0	0	6	14	8	2	6	13	7	4	3	5	1	14	7	6	5	19
4	12	73	100	62	0	0	0	247	0	92	80	53	22	23	79	60	85	12	40	30	112	53	63	43	141
5	2044	2660	902	1116	0	0	0	0	6722	1444	2355	1897	1026	1385	2687	1249	1401	3526	1016	261	1085	834	4842	849	1031
1	336	582	277	369	16	6	6	92	1444	1564	0	0	0	46	276	460	782	782	238	68	272	204	1088	204	272
2	693	942	400	449	28	7	14	80	2355	0	2484	0	0	322	1150	552	460	1242	378	108	432	324	1728	324	432
3	670	799	232	277	14	6	8	53	1897	0	0	1978	0	598	874	276	230	989	301	86	344	258	1376	258	344
4	357	433	146	122	7	1	2	22	1026	0	0	0	1058	460	506	46	46	529	161	46	184	138	736	138	184
1	480	601	202	143	10	2	6	23	1385	46	322	598	460	1426	0	0	0	713	217	62	248	186	992	186	248
2	908	1107	371	420	20	7	13	79	2687	276	1150	874	506	0	2806	0	0	1403	427	122	488	366	1952	366	488
3	403	495	199	237	15	3	7	60	1249	460	552	276	46	0	0	1334	0	667	203	58	232	174	928	174	232
4	265	553	283	417	20	8	4	85	1401	782	460	230	46	0	0	0	1518	759	231	66	264	198	1056	198	264
1	1746	1505	215	76	1	0	3	12	3526	782	1242	989	529	713	1403	667	759	3542	0	0	0	0	3542	0	0
2	98	370	202	408	13	4	5	40	1016	238	378	301	161	217	427	203	231	0	1078	0	0	0	308	616	154
3	21	115	100	72	11	5	1	30	261	68	108	86	46	62	122	58	66	0	0	308	0	0	154	0	154
4	128	527	315	262	14	7	14	112	1085	272	432	344	184	248	488	232	264	0	0	0	1232	0	616	0	616
5	63	239	223	399	26	4	7	53	834	204	324	258	138	186	366	174	198	0	0	0	0	924	308	308	308
1	1934	2083	547	364	15	2	6	63	4842	1088	1728	1376	736	992	1952	928	1056	3542	308	154	616	308	4928	0	0
2	36	205	175	508	26	1	5	43	849	204	324	258	138	186	366	174	198	0	616	0	0	308	0	924	0
3	86	468	333	345	24	17	19	141	1031	272	432	344	184	248	488	232	264	0	154	154	616	308	0	0	1232
	1	2	3	4	1	2	3	4	5	1	2	3	4	1	2	3	4	1	2	3	4	5	1	2	3