

**ANTONIA SILVA PAREDES MOREIRA
BRIGIDO VIZEU CAMARGO
JORGE CORREIA JESUINO
SHEVA MAIA DA NÓBREGA
(ORGANIZADORES)**

**PERSPECTIVAS
TEÓRICO-METODOLÓGICAS
EM REPRESENTAÇÕES SOCIAIS**

**Prefácio
SERGE MOSCOVICI**

**EDITORA UNIVERSITÁRIA - UFPB
JOÃO PESSOA, PARAÍBA, BRASIL.
2005**

CAPÍTULO 17

ALCESTE: UM PROGRAMA INFORMÁTICO DE ANÁLISE QUANTITATIVA DE DADOS TEXTUAIS.

BRÍGIDO VIZEU CAMARGO

Introdução

A análise de dados textuais, parte importante da análise de conteúdo, utiliza a linguagem escrita ou transcrita (material escrito da mídia, falas durante uma entrevista, depoimentos, relatos, documentos, etc.). No Brasil, no campo das Ciências Humanas, os textos são considerados um material de pesquisa que preserva a qualidade dos fenômenos estudados, pois são produzidos de forma “mais natural” que as respostas aos instrumentos padronizados, como questionários estandardizados ou testes. Mas raramente o dado textual fornece material natural, ou seja, um produto de interações sociais independentes do quadro da pesquisa que os utiliza. Isto ocorre somente, e mesmo assim parcialmente, no caso da pesquisa documental. Em outras palavras, segundo Henry e Moscovici (1968), uma análise de conteúdo tem dois planos: o vertical (análise das condições de produção) e o horizontal (análise dos textos), o procedimento da análise horizontal é determinado pelo plano vertical (objetivos da análise, escolha do material) e este último também é determinado pelos textos, na medida em que o textual deve ter as condições da sua produção consideradas.

Outro problema, decorrente do problema apontado no parágrafo precedente, é acreditar que quando se trabalha com dados textuais pode-se abandonar as exigências de uma análise sistemática e objetiva. Que, para isto, basta anunciar que estamos realizando uma “pesquisa qualitativa”, e que a simples transcrição dos extratos dos textos coletados têm um poder mágico de autorizar interpretações sobre o que se estuda.

A análise quantitativa de dados textuais não deixa de considerar a qualidade do fenômeno estudado, e ainda fornece critérios provenientes do próprio material para a consideração do mesmo como indicador de um fenômeno de interesse científico. Uma análise quantitativa de textos transcritos ou escritos tem como base as leis de distribuição dos seus

respectivos vocabulários. Durante a década de 70, os trabalhos de lexicografia contribuíram para este tipo de análise. E nos últimos 20 anos, a aplicação da informática permitiu o desenvolvimento deste tipo de análise. Bardin (1991) apresenta uma boa revisão dos trabalhos pioneiros no uso da informática em análise de conteúdo.

Na França, existem vários programas informáticos para análise textual. Os principais são: o ALCESTE de M. Reinert, o Tri-Deux (MOTS) de P. Cibois, o SPAD-T de L. Lebart, o LEXICO de A. Salem e o Hyperbase de E. Brunet (Ver Reinert, 1990; Cibois, 1989 e 1990; Doise, Clemence e Lorenzi-Cioldi, 1992; Lebart e Salem, 1994).

O programa informático ALCESTE

O objetivo deste trabalho é descrever o programa informático ALCESTE[©]? (Análise Lexical Contextual de um Conjunto de Segmentos de Texto)¹ e instrumentalizar os pesquisadores brasileiros, que não estão habituados a língua francesa e as análises textuais de segmentos de texto, a utilizá-lo.

Em 1998, o programa informático ALCESTE foi introduzido no Brasil (Veloz, Nascimento-Schulze e Camargo, 1999). Ele apresenta um interesse particular, pois empregando uma análise de classificação hierárquica descendente, além de permitir uma análise lexicográfica do material textual, oferece contextos (classes lexicais) que são caracterizados pelo seu vocabulário e pelos segmentos de textos que compartilham este vocabulário (IMAGE, 1998).

As noções de *corpus*, “unidade de contexto inicial” e “unidade de contexto elementar”

O programa ALCESTE toma como base um único arquivo, mas deve-se indicar o que Reinert (1990) chama de unidades de contexto iniciais (UCIs) e preparar este arquivo segundo certas regras (Ver anexo A: Regras para a preparação do material textual).

©Todos os direitos reservados à Société IMAGE (1986-2000). Autor: Max Reinert (Licença CNRS/UTM).

¹ “Analyse Lexicale par Contexte d’ un Ensemble de Segments de Texte”.

A definição destas unidades é feita pelo pesquisador e depende da natureza da pesquisa. Se a análise vai ser aplicada a um conjunto de entrevistas, cada uma delas será uma UCI. Caso a análise diga respeito às respostas de “n” participantes a uma questão aberta, cada resposta será uma UCI e teremos “n” unidades de contexto iniciais (UCIs). Quando se tratar de artigos de jornais, atas de reuniões, cartas, etc; cada exemplar destes documentos será uma UCI.

Um conjunto de unidades de contexto iniciais (UCIs) constitui um *corpus* de análise. O *corpus* adequado à análise do ALCESTE deve constituir-se num conjunto textual centrado em um tema. O material textual tem que ser monotemático, pois a análise de textos sobre vários itens previamente estruturados ou diversos temas resulta na reprodução da estruturação prévia dos mesmos.

No caso de entrevistas, onde há falas que produzem textos mais extensos, desde que o grupo seja homogêneo, é suficiente entre 20 e 30 UCIs (Ver Ghiglione e Matalon, 1993, p. 60). Se o delineamento é comparativo, sugere-se pelo menos 20 UCIs para cada grupo.

Em se tratando de respostas a questões abertas de um questionário, cada UCI vai ser composta da adição dos textos obtidos das respostas somente quando elas se referirem a um mesmo tema. Caso as questões referiram-se a temas ou aspectos diferentes, é necessário realizar uma análise para cada questão. Como mencionado anteriormente, a análise feita pelo programa ALCESTE é sensível à estruturação do estímulo que produz o material textual, e isto é uma importante fonte de invalidação das conclusões.

Quando as respostas apresentarem uma média em torno de três ou quatro linhas, é necessário um número bem maior de respostas para a constituição de um *corpus* de análise. Para um funcionamento ótimo, o ALCESTE deve analisar *corpus* com um tamanho em torno de 1000 linhas. No caso da linha ter 70 caracteres o texto teria 70.000, o que significa aproximadamente 20 páginas de texto em fonte Courier 10, com espaçamento simples. Isto significa que há necessidade de aproximadamente 250 respostas curtas. No entanto, foram realizadas análises sobre *corpus* com metade do tamanho recomendado e com sucesso. A própria análise utilizada, a seguir, neste trabalho é um exemplo disto.

As UCI são separadas por linhas de comando também chamadas de “linhas com asteriscos”. No caso de entrevistas, por

exemplo, como cada uma delas é uma unidade de contexto inicial (UCI), elas necessariamente devem começar com uma linha de comando. Esta linha informa o número de identificação do entrevistado (do produtor do texto que se segue) e algumas características (variáveis) que são importantes para o delineamento da pesquisa (como sexo, faixa etária, afiliação a determinados grupos, nível social e cultural, etc.). Isto depende de cada pesquisa e o número de modalidades de cada uma destas variáveis depende do delineamento da pesquisa e do número de entrevistas realizadas.

Embora seja o pesquisador que define as unidades de contexto iniciais (UCIs), nem sempre é ele que controla a divisão do corpus em segmentos de texto, ou seja nas unidades de contexto elementar (UCEs). Numa análise padrão, após reconhecer as indicações das UCIs é o programa ALCESTE divide o material em unidades de contexto elementar (UCEs). Elas são segmentos de texto, na maior parte das vezes, do tamanho de três linhas, dimensionadas pelo programa informático em função do tamanho do corpus e, em geral, respeitando a pontuação. Mas o pesquisador pode configurar a divisão das UCEs, e no caso de uma grande quantidade de respostas curtas a uma pergunta aberta de um questionário, aconselha-se que as unidades de contexto elementar (UCEs) sejam definidas enquanto unidades de contexto inicial (UCIs), ou seja, enquanto a resposta dada à questão.

Etapas de análise do ALCESTE

Para iniciar uma análise utilizando o programa informático ALCESTE consultar o anexo B: Guia de operações para o lançamento e a utilização do programa ALCESTE. Iniciada a análise, o programa executa 4 etapas (A, B, C e D), cada uma contendo três operações, exceto a última (D) que contém cinco.

Etapas A: Leitura do texto e cálculo dos dicionários

Nesta etapa o programa prepara o corpus, reconhece as UCIs, faz uma primeira segmentação do texto, agrupa as ocorrências das palavras em função de suas raízes e procede ao cálculo da frequência destas formas reduzidas.

Operações:

A1- Re-formatação e divisão do texto em segmentos de tamanho similar (UCEs).

A2- Pesquisa do vocabulário e redução das palavras com base em suas raízes (formas reduzidas) Por exemplo: as palavras filho, filha, filhos e filhas, são computadas sob a forma reduzida filh+.

A3- Criação do dicionário de formas reduzidas.

Etapa B: Cálculo das matrizes de dados e classificação das UCEs

Esta é uma etapa de cálculo. As UCEs são classificadas em função dos seus respectivos vocabulários, e o conjunto delas é repartido em função da frequência das formas reduzidas. A partir de matrizes cruzando formas reduzidas e UCEs, variando o tamanho das UCEs, aplica-se o método de classificação hierárquica descendente (CHD) e obtém-se uma classificação definitiva. Esta análise visa obter classes de UCEs que, ao mesmo tempo, apresentam vocabulário semelhante entre si, e vocabulário diferente das UCEs das outras classes. Para isto, utiliza-se do teste do *qui-quadrado* de associação das formas reduzidas e das UCEs às classes.

Operações:

B1- Seleção das UCEs a serem consideradas e cálculo da matriz: formas reduzidas x UCEs.

B2- Cálculo das matrizes de dados para Classificação Hierárquica Descendente (CHD).

B3- Classificação Hierárquica Descendente.

Etapa C: Descrição das classes de UCEs

Esta etapa fornece os resultados mais importantes. Aqui o programa apresenta o dendograma da classificação hierárquica descendente (CHD), que ilustra as relações entre as classes. O programa executa cálculos complementares para cada uma das classes, obtidas pelos cálculos da etapa precedente, e fornece resultados que nos permite a descrição de cada uma das classes, principalmente, pelo seu vocabulário característico (léxico) e pelas suas palavras com asterisco (variáveis). Além disto, o programa fornece uma outra forma de apresentação dos resultados, através de uma análise fatorial de correspondência feita a partir da CHD.

Operações:

C1- Definição das classes escolhidas.

C2- Descrição das classes.

C3- Análise Fatorial de Correspondência ou AFC (representação das relações entre as classes num plano fatorial).

Etapa D: Cálculos complementares

A Etapa D é um prolongamento da Etapa C. Com base nas classes de UCEs escolhidas, o programa calcula e fornece-nos as UCEs mais características de cada classe, permitindo a contextualização do vocabulário típico de cada classe obtido na operação C2. Ele fornece ainda Classificações Hierárquicas Ascendentes (CHAs) para cada classe, permitindo assim o estudo das relações dos elementos (formas) intraclasse, os dendogramas constituem-se bons guias para a análise das relações das palavras de cada classe obtidas na operação C2. Outros recursos são possíveis, tais como o tratamento de segmentos repetidos nas UCEs e a exportação destas UCEs para outros programas informáticos.

Operações:

D1- Seleção das UCE mais características de cada classe.

D2- Pesquisa dos segmentos repetidos (duplas e segmentos) por classe.

D3- Classificação Hierárquica Ascendente (CHA). Realizada a partir da construção de uma matriz de formas associadas a uma mesma classe x UCEs da classe associada. Com isto chega-se a representações arborescentes destas relações.

D4- Seleção das palavras mais características das classes para apresentação em um *index* de contexto de ocorrência.

D5- Exportação para outros programas de *sub-corpus* de UCEs por classe.

Em resumo, as operações onde se produzem os resultados mais importantes para a interpretação de um corpus são a C1 (dendograma da classificação hierárquica descendente - CHD), C2 (descrição das classes), D1 (seleção das UCE mais características de cada classe) e D3 (classificação hierárquica ascendente das palavras por classe - CHA). A operação A2, onde há a produção do dicionário das formas reduzidas (arquivo: "A2_dico"), também tem a sua importância, pois sempre se pode melhorar este processo de redução das palavras as suas raízes, o que aprimorará globalmente a análise do corpus. Para isto se utiliza o procedimento de uma nova análise com o uso de um dicionário corrigido, denominado: "Dicin" (Ver o Anexo B).

O significado das classes obtidas pelo ALCESTE

O que são estas classes de palavras e de unidades de contexto elementar? Em nível do programa informático, cada classe é composta de várias UCEs em função de uma classificação segundo a distribuição do vocabulário destas UCEs. Em nível interpretativo Reinert (1990) utiliza a noção de “mundo”, enquanto um quadro perceptivo-cognitivo com certa estabilidade temporal associado a um ambiente complexo. Inicialmente ele aplicou à obra literária o programa ALCESTE. Em pesquisas no campo da lingüística estas classes são interpretadas como campos lexicais (Cros, 1993) ou contextos semânticos. Em pesquisas no campo da psicologia social, particularmente aquelas interessadas em estudar o conhecimento do senso comum, tendo em vista o estatuto que elas conferem às manifestações lingüísticas, estas classes podem indicar representações sociais ou campos de imagens sobre um dado objeto, ou somente aspectos de uma mesma representação social (Veloz, Nascimento-Schulze e Camargo, 1999). Na maior parte das vezes não há coincidência entre o número de classes e o número de representações sociais envolvidas, como ocorreu no estudo relativo à referência que se acabou de citar. O que vai definir se elas indicam representações sociais ou apenas uma representação social é o seu conteúdo, e sua relação com fatores ligados ao plano geral de cada pesquisa, geralmente expresso na seleção diferenciada dos participantes segundo sua afiliação grupal, suas práticas sociais anteriores, etc.

Uma aplicação do ALCESTE: Análise de entrevistas com “Garotas de Programa” sobre a prevenção da gravidez e das doenças sexualmente transmissíveis

Para ilustrar a aplicação do programa informático ALCESTE, vai-se relatar, em linhas gerais, uma análise que o utilizou. Trata-se de parte de uma dissertação de mestrado (Oltramari, 2001). Foram entrevistadas 40 mulheres, profissionais do sexo, caracterizadas por cinco variáveis: faixa de idade, existência de parceiro sexual estável e não comercial, se têm filho ou não, há quanto tempo pratica sexo venal e tipos de atribuição de causa para fazer isto. Estas garotas apresentavam idade mínima de 19 e máxima de 47 anos. Das entrevistadas, 18 declararam ter relações sexuais estáveis e não comerciais com um mesmo parceiro e 22 não preencheram esta condição (ter parceiro sexual

fixo). O *corpus* analisado foi denominado “prevenir”. Ele foi composto das respostas das entrevistadas sobre as formas de prevenção da gravidez e das doenças sexualmente transmissíveis (DSTs). Mais precisamente, as perguntas foram as seguintes:

- Você utiliza algum método para evitar gravidez?
- Qual método você mais utiliza?
- Porque utiliza este método?
- O que são doenças sexualmente transmissíveis para você?
- Você alguma vez pegou alguma doença destas?
- Como você faz para se prevenir das doenças sexualmente transmissíveis?

O anexo A contém um extrato deste *corpus* e a estrutura das suas linhas de comando, que separaram as respostas de cada garota profissional do sexo. A seguir, fornecem-se elementos do relatório completo (*rapport complét*), para que se possa fazer um exame sintético do mesmo.

Etapa A - Leitura do texto e cálculo dos dicionários

Operação A1- Leitura do corpus (*Lecture du corpus*)

Número de linhas com estrelas (*Nombre de lignes étoilées*) = 40 (o programa reconhece a separação do corpus em 40 unidades de contexto inicial – UCIs - ou entrevistas).

Operação A2 – Cálculo do dicionário (*Calcul du dictionnaire*)

Nº de formas distintas ou palavras diferentes (*Nombre de formes distinctes*) = 1.084.

Nº de ocorrências (*Nombre d’occurrences*) = 6.491.

Frequência média por forma distinta (*Fréquence moyenne par forme*) = 6 (o produto da divisão entre o Nº de ocorrências e o Nº de formas distintas).

Nº de palavras com frequência igual a um (*Nombre de hapax*) = 581 (um valor alto indica a heterogeneidade do vocabulário empregado pelos que produziram os textos que compõem o *corpus*).

Etapa B - Cálculo das matrizes de dados e classificação das UCEs

Operação B1- Seleção das UCEs e cálculo dos dados (*Sélection des uce et calcul des données*)

B11: Frequência mínima de uma palavra analisada (*Fréquence minimum d'un mot analysé*) = 4.

Nº de palavras analisadas (*Nombre de mots analysés*) = 158 (com frequência e" 4).

Nº de palavras instrumento (*Nombre de mots supplémentaires de type "r"*) = 107.

Nº de palavras com asteriscos ou palavras variáveis = 54 (*Nombre de mots supplémentaires de type "s"*).

Nº de linhas da matriz de dados ou total de palavras (*Nombre de lignes de B1_DICB*) = 319.

Nº de ocorrências analisadas (*Nombre d'occurrences analysées*) = 2.229.

Após a redução das palavras as suas raízes, obtiveram-se 158 palavras analisáveis (com frequência igual ou superior a quatro), 107 palavras instrumento e 54 palavras variáveis (com asterisco). As 158 palavras analisáveis ocorreram 2.229 vezes.

Etapa C - Descrição das classes de UCEs

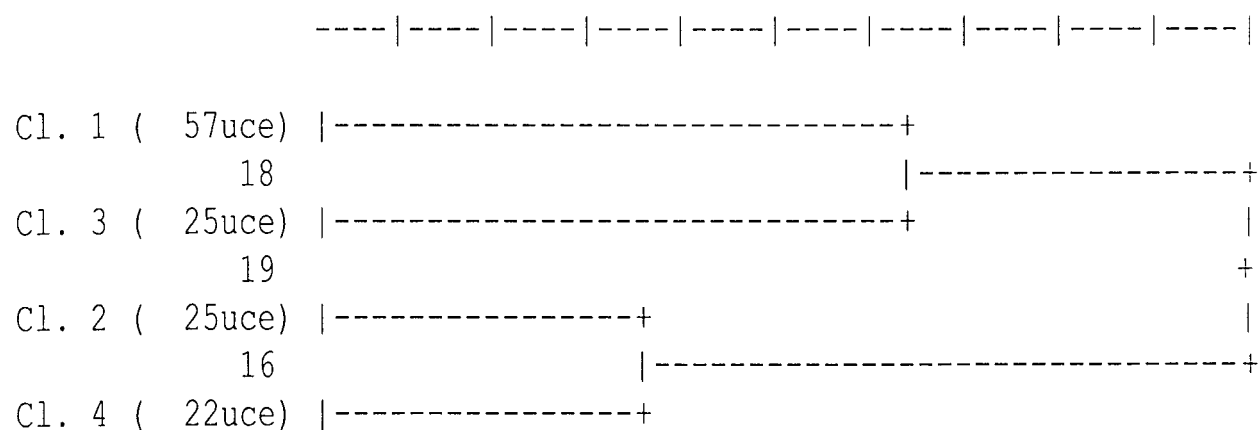
Operação C1- Intersecção entre as Classes (*Intersection des classes*)

129 UCEs classificadas de 163 selecionadas, ou seja 79,14% (*129 u.c.e classées sur 163 soit 79.14 %*)

Este corpus foi dividido em 163 UCEs das quais 129, ou seja, 79, 14% do total, foram designadas numa mesma classe em duas classificações hierárquica descendentes que utilizaram UCEs com tamanhos diferentes.

Classificação Hierárquica Descendente (CHD) (*Classification Descendante Hiérarchique -CDH*).

Dendograma das Classes Estáveis (*Dendrogramme des classes stables*) :



Este dendograma, que representa o produto da CHD, ilustra as relações interclasses. Lê-se da direita para a esquerda. Num primeiro momento, o *corpus* “prevenir” foi dividido (1ª partição) em dois sub-corpus (de um lado aquele que vai originar as classes 1 e 3, e do outro o referente às classes 2 e 4). Num segundo momento o primeiro sub-corpus foi dividido em dois (2ª partição), assim obteve-se as classes 1 e 3. E num terceiro momento, o segundo sub-corpus foi dividido em dois (3ª partição), resultando as classes 2 e 4. A CHD parou aqui, pois as 4 classes mostraram-se estáveis, ou seja: compostas de unidades de contexto elementar (UCEs) com vocabulário semelhante.

Operação C2 - Descrição das Classes (léxica) (*Profil des classes*)

Nº de classes (*Nombre de Classes*) = 4.

Para a análise descritiva do vocabulário de cada classe, uma possibilidade é utilizar dois critérios simultâneos: 1) reter a atenção nas palavras não instrumentais com média maior que 6 (critério lexicográfico indicado na operação A2) e 2) considerar aquelas com c^2 de associação à classe ³ 3,84 (pois o cálculo deste teste estatístico é feito com base em uma tabela com grau de liberdade igual a 1). Nesta operação o programa oferece uma lista de formas reduzidas associadas a cada classe. A seguir, por uma questão de espaço, apresenta-se a lista abreviada referente à classe 1.

Classe N° 1 (Contexto A)

Número de UCES (*Nombre d'u.c.e.*) : 57 ou seja (“soit”):

44.19 %

| N° | freq. na classe e no corpus | | % na classe | χ^2 | forma reduzida (palavra) |
|-------|-----------------------------|-----|-------------|----------|--------------------------|
| num | effectifs | | pourc. | chi2 | identification |
| 7 | 4. | 5. | 80.00 | 2.71 | ao |
| 8 | 22. | 30. | 73.33 | 13.47 | as |
| 18 | 7. | 9. | 77.78 | 4.43 | comec+ |
| 25 | 13. | 22. | 59.09 | 2.39 | da |
| 31 | 7. | 9. | 77.78 | 4.43 | deu+ |
| 33 | 5. | 6. | 83.33 | 3.91 | dia+ |
| 36 | 7. | 7. | 100.00 | 9.35 | dinheiro |
| 44 | 5. | 7. | 71.43 | 2.23 | entr+ |
| 46 | 3. | 3. | 100.00 | 3.88 | esclarecimento+ |
| 49 | 5. | 5. | 100.00 | 6.57 | exam+ |
| 58 | 17. | 27. | 62.96 | 4.88 | faz+ |
| 59 | 8. | 12. | 66.67 | 2.71 | fic+ |
| 62 | 5. | 6. | 83.33 | 3.91 | fiquei |
| (...) | | | | | |
| 96 | 5. | 5. | 100.00 | 6.57 | olh+ |
| 97 | 4. | 4. | 100.00 | 5.21 | oral |
| 99 | 6. | 7. | 85.71 | 5.18 | palestras |
| 101 | 6. | 6. | 100.00 | 7.95 | pass+ |
| 109 | 13. | 20. | 65.00 | 4.16 | peesso+ |
| 114 | 3. | 3. | 100.00 | 3.88 | prefiro |
| 124 | 3. | 3. | 100.00 | 3.88 | questao |
| (...) | | | | | |
| 158 | 7. | 8. | 87.50 | 6.49 | vou |
| 161 * | 41. | 68. | 60.29 | 15.13 * | o |
| (...) | | | | | |
| 271 * | 46. | 91. | 50.55 | 5.07 * | *fil_1 |
| 274 * | 41. | 78. | 52.56 | 5.62 * | *ida_2 |
| 283 * | 5. | 6. | 83.33 | 3.91 * | *ind_09 |
| 284 * | 4. | 5. | 80.00 | 2.71 * | *ind_10 |
| (...) | | | | | |
| 315 * | 33. | 60. | 55.00 | 5.32 * | *par_1 |
| 319 * | 28. | 52. | 53.85 | 3.30 * | *temp_3 |

As palavras analisáveis grifadas podem ser consideradas os elementos mais importantes, conforme o duplo critério já explicitado, para descrever esta classe. As palavras com asterisco, no final da tabela, indicam que os indivíduos que compartilham a modalidade das variáveis que elas representam foram os principais responsáveis pela produção das UCES da classe em foco. Sendo assim, a maior parte das UCES da classe 1 foi produzida por garotas de programa mais velhas (*fil_2, de 27 a 47 anos), que têm um parceiro sexual fixo (*par_1) e que têm filho (*fil_1).

Numa publicação, geralmente, a descrição das classes é apresentada sob a forma de tabela (Ver Tabela I).

Tabela I - Palavras associadas significativamente à classe 1 (57 UCEs - 44.19% do total)

| Palavra | Frequência | χ^2 |
|-----------|------------|----------|
| Pessoas | 13 | 4,16 |
| Deus | 7 | 4,43 |
| Dinheiro | 7 | 9,35 |
| Palestras | 6 | 5,18 |
| Fazer | 17 | 7,88 |
| Começar | 7 | 4,43 |

Em função do número de UCEs que compõem esta classe (57, ou 44,19% do total das UCEs classificadas), ela é a que caracteriza, em maior proporção, o conteúdo do *corpus* “prevenir”. Esta operação nos fornece uma descrição em nível das palavras, das noções contempladas pela significação.

Esta classe organiza-se em torno de quatro elementos: *pessoas, Deus, dinheiro e palestras*. É necessário lembrar que estes elementos foram utilizados nas respostas sobre prevenção de doenças sexualmente transmissíveis e contracepção. E para compreender o conteúdo desta classe, é necessário situar seus principais elementos (palavras) nas UCEs significativamente associadas à ela.

Operação D1 - UCEs significativas de cada classe (*Tri des uce par classe*)

O programa ALCESTE fornece também uma lista de unidades de contexto elementar (UCEs) típicas de cada classe. Cada UCE é precedida do seu número de ordem no corpus e de um índice que indica o coeficiente de associação dela a classe em questão. Este coeficiente é o mesmo utilizado para as palavras, ou seja: o *qui-quadrado*²

² Lista dos índices e seu valor em χ^2 (*Liste des valeurs de clé*) :

0 se $\chi^2 < 2.71$ 3 se $\chi^2 < 6.63$ 6 se $\chi^2 < 30.00$ acima de 8 se $\chi^2 > 50.00$
 1 se $\chi^2 < 3.84$ 4 se $\chi^2 < 10.80$ 7 se $\chi^2 < 40.00$

2 se $\chi^2 < 5.02$ 5 se $\chi^2 < 20.00$ 8 se $\chi^2 < 50.00$

Têm-se, a seguir, extratos da lista de UCES selecionadas da classe 1 (ou contexto lexical A).

Clé sélectionnée : A

(...)

92 10 O #pessoal do arco_iris distribuem folhetos e a #gente le, entao #da para #saber identificar. #as dst saem para-fora, se eu vejo que o negocio esta #meio esquisito, ja devolvo o #dinheiro e mando embora.

(...)

162 10 com eu tive coragem, #as #vezes o #dinheiro #fala mais alto, no meu caso muitas #vezes o #dinheiro #falou mais alto, foi #deus que me protegeu e hoje eu nao tenho nada, contei muito com a #sorte, porque eu realmente vacilei um monte.

(...)

155 8 apesar dos #homens nao querer preservativo de-jeito-nenhum, a maioria deles nao querem, se eu fosse atender #oral sem preservativo, eu atenderia de dez a quinze #pessoas por #dia, eles #deixam de #fazer o #programa por causa do preservativo, de #fazer #oral.

(...)

51 7 #as #vezes o #dinheiro que voce ganha num #programa, nao paga o #tratamento que voce tem que #fazer, isso, se voce tiver #sorte e nao pegar uma aids. se voce pegar, seu futuro estara perdido, psicologico, emocional, financeira e voce #comeca a se abalar, entao a #gente tem muita preocupacao com isso.

63 7 infelizmente, #fiquei sabendo hoje, pequei o ultra som hoje, mas la para frente eu #vou #ficar. eu conheco bastante doencas sexualmente transmissiveis, porque eu trabalhei em uma #epoca em foz de iguacu, uma #vez por #mes ia um #pessoas de uma ong tipo o arco iris, esse #pessoal dava #palestras, tinha telao, (...)

O foco desta classe é a prevenção das doenças sexualmente transmissíveis (DSTs). O verbo *fazer* refere-se a *programas* (ter relações sexuais por dinheiro). A palavra *dinheiro* é central para as mulheres mais velhas compreenderem o risco que correm nesta profissão.

Quando a garota arrisca, ou seja, quando ela tem relações sexuais sem preservativo por mais dinheiro e não contrai doenças, este fato é atribuído à proteção de *Deus*. O termo *pessoas* indica principalmente os clientes. O *dinheiro* aparece, várias vezes, como instrumento de imposição dos clientes para que a profissional faça coisas arriscadas para a sua saúde e sua vida. O conhecimento do risco para a saúde, ou seja, das DSTs, está associado à palavra *palestras*, feitas por pessoal de Organizações não-Governamentais. As noções ou palavras desta classe estão pouco articuladas. Mas, em resumo, aqui aparece a idéia de que o perigo para a saúde advém da ganância, de fazer tudo para ganhar dinheiro.

Classe 3 (Contexto C)

Vizinha da classe precedente (Ver Dendograma da CHD), a maior parte das UCEs da classe 3 foram produzidas por garotas de programa que não têm um parceiro sexual fixo (*par_2). A tabela II apresenta as principais palavras que a caracteriza.

Tabela II- Palavras associadas significativamente à classe 3 (25 UCEs - 19.38% do total)

| Palavra | Frequência | χ^2 |
|---------|------------|----------|
| Coisas | 9 | 12,55 |
| Mulher | 4 | 6,76 |
| Protejo | 14 | 49,70 |
| Usando | 12 | 44,23 |
| Sei | 11 | 5,53 |
| Peguei | 11 | 40,74 |
| Pegar | 9 | 47,37 |

Duas noções se destacam nesta classe: *coisas* e *mulher*. Aqui os verbos são mais elucidativos do que estas duas noções. A dupla *usar* (camisinha) e não *pegar* (DSTs e AIDS) indica a compreensão do que é se prevenir das DSTs.

Extratos da lista de UCEs típicas, ou associadas significativamente a esta classe 3, são apresentado a seguir.

Clé sélectionnée : C

(...)

148 15 eu nunca #peguei dst nao, so #corrimento #normal que toda a #mulher tem e geralmente fica mais forte quando voce esta gravida. eu ja tive #corrimento forte mas nao chegou a ser a #uma dst, e #uma #coisa natural da #mulher mesmo. eu me #protejo das dst #usando camisinha.

(...)

82 11 O que eu sei e isso e o resto e vendo, de #nome e dificil gravar, sao varios e o #hiv o terror de todas, que bate todas. eu me #protejo #usando camisinha sem-duvida, eu #acho que #uma ainda e ainda pouco, tem que ser duas.

124 8 porque tem #mulher que #vai sem camisinha, porque, isso voce #pode apostar de cem #pode se tirar #uma que #diga que nunca foi sem camisinha. eu nao sei te #explicar muito sobre dst, mas eu ja li muito sobre isso, pelo beijo voce #pode #pegar qualquer doenca transmissivel, ou qualquer ato #sexual, dependendo, voce sabe que tem.

(...)

Conforme estas UCEs a proteção aparece diretamente ligada ao uso de preservativo. Mas o uso do preservativo nem sempre é bem conhecido, usar duas camisinhas, além de não ser funcional, pode, ao contrário, não proteger os protagonistas da relação sexual com penetração. A palavra *coisas* não aparece só conotando corrimento vaginal, como uma das UCEs ilustra, mas também conota DST.

Esta classe, típica do pensamento de garotas que não tem parceiro fixo, diferentemente da classe 1 (onde as garotas têm parceiro), apresenta uma concepção de prevenção das DST que inclui o uso do preservativo. Isto se inscreve no contexto do contato com os clientes, que é considerado como potencialmente perigoso, já que pode transmitir este tipo de doença. As UCEs associadas a esta classe indicam uma concepção da prevenção de DST que tem como cerne evitar, ao máximo, o contato. Isto pode parecer paradoxal, em se tratando de profissionais do sexo, já que a relação sexual é o contato corporal mais estreito entre as várias formas de contatos interpessoais.

Classe 2 (Contexto B)

Segundo o dendograma da CHD, esta classe, juntamente com a classe 4, se opõe as duas classes já analisadas. Esta diferenciação parece ter origem na forma como foi estruturada a entrevista. Por um lado, procurou-se saber o que as garotas de programa sabiam e pensavam sobre a prevenção de DST, e por outro lado perguntou-se também a mesma coisa sobre a contracepção.

Tabela III- Palavras associadas significativamente à classe 2 (25 UCEs - 19.38% do total)

| Palavra | Frequência | χ^2 |
|-----------|------------|----------|
| Gravidez | 19 | 30,80 |
| Camisinha | 18 | 14,43 |
| Doenças | 16 | 11,57 |
| Dst | 15 | 9,92 |
| Métodos | 13 | 22,84 |
| Aids | 9 | 11,17 |
| Gonorréia | 7 | 5,95 |
| Noite | 5 | 6,51 |
| Namorado | 4 | 3,89 |
| Utilizo | 25 | 26,20 |
| Evitar | 19 | 33,94 |
| Sei | 15 | 19,30 |

Conforme a tabela III, a concepção mais compartilhada entre as garotas de programa sobre a prevenção da gravidez (pois nenhuma característica delas foi associada a esta classe) organiza-se, principalmente, em torno dos termos: *gravidez, camisinha, doenças (DSTs, AIDS, Gonorréia) e métodos.*

Clé sélectionnée: B

147 20 eu nao utilizo nenhum #metodo para #evitar a #gravidez. eu #sei que as #dst se transmitem se voce nao usar #camisinha. eu #sei que tem a #sifilis, a #gonorreia, #aids, tem varios #tipos de #doenca transmissiveis, com #os meus #parceiros eu #uso #camisinha direto na #noite, mas com o meu #namorado realmente eu nao #uso.

(...)

96 10 O que eu #sei sobre #dst, e que ha varios #tipos de #doencas, hepatite, #sifilis, #gonorreia, #aids e outras mais. eu ja peguei #gonorreia, a anos atras. quando eu comecei, nao se falava tanto em #aids e muito menos em #camisinha, a #camisinha era um #metodo para #evitar #gravidez, apenas, isso digamos a dezesseis anos atras, nao se falava, estava comecando a aparecer a #aids.

(...)

144 5 para #evitar a #gravidez eu utilizo #camisinha e anticoncepcional. eu utilizo #os dois porque e mais seguro. eu #sei praticamente tudo sobre #dst, #aids, tem varias #doencas, #gonorreia. uma vez eu peguei, condiloma, foi do meu #ex_marido. eu fui casada antes-de trabalhar na #noite. eu utilizo a #camisinha para me #proteger #das #dst.

Os extratos, que contém três UCEs significativas para a classe 2, indicam que *gravidez* e as *DSTs* são dois problemas importantes a serem enfrentados pelas profissionais do sexo. No entanto, o primeiro não é uma doença e o segundo sim. Aqui o preservativo aparece tanto como método de prevenção das DSTs como da gravidez. No entanto o uso do preservativo foi mais associado às relações sexuais venais do que as relações com um parceiro fixo (onde, provavelmente, há maior envolvimento emocional). Mesmo assim, no sexo venal, quando se oferece mais dinheiro o uso do preservativo pode não ser observado (como mostra o conteúdo da classe 1).

Quando temos uma maior quantidade de noções em uma classe, e elas se apresentam estruturadas, como é o caso desta classe, a operação D3 fornece indicações úteis para compreendermos a organização interna dos seus elementos.

Operação D3 – Classificação Hierárquica Ascendente das palavras por classe (C.A.H. des mots par classe)

A operação D3 fornece indicações úteis para se entender as relações entre as palavras características de cada classe. O dendograma apresentado a seguir ilustra as relações dos principais elementos da classe 2 (ou contexto lexical B).

