

## Guia de Utilização do Software Alceste: uma ferramenta de análise lexical aplicada à interpretação de discursos de atores na agricultura



ISSN 1517-5111  
ISSN online 2176-5081  
Outubro, 2009

*Empresa Brasileira de Pesquisa Agropecuária  
Embrapa Cerrados  
Ministério da Agricultura, Pecuária e Abastecimento*

## ***Documentos 275***

### **Guia de Utilização do Software Alceste: uma ferramenta de análise lexical aplicada à interpretação de discursos de atores na agricultura**

*Evie dos Santos de Sousa  
Márcio Alexandre dos Santos Rodrigues  
Francisco Eduardo de Castro Rocha  
Cíntia Ribeiro Martins*

Embrapa Cerrados  
Planaltina, DF  
2009

Exemplares desta publicação podem ser adquiridos na:

**Embrapa Cerrados**

BR 020, Km 18, Rod. Brasília/Fortaleza  
Caixa Postal 08223  
CEP 73310-970 Planaltina, DF  
Fone: (61) 3388-9898  
Fax: (61) 3388-9879  
<http://www.cpac.embrapa.br>  
[sac@cpac.embrapa.br](mailto:sac@cpac.embrapa.br)

**Comitê de Publicações da Unidade**

Presidente: *Fernando Antônio Macena da Silva*  
Secretária-Executiva: *Marina de Fátima Vilela*  
Secretária: *Maria Edilva Nogueira*

Supervisão editorial: *Jussara Flores de Oliveira Arbués*  
Equipe de revisão: *Francisca Elijani do Nascimento*  
*Jussara Flores de Oliveira Arbués*  
Assistente de revisão: *Elizelva de Carvalho Menezes*  
Normalização bibliográfica: *Shirley da Luz Soares Araújo*  
Editoração eletrônica: *Leila Sandra Gomes Alencar*  
Capa: *Leila Sandra Gomes Alencar*

Impressão e acabamento: *Alexandre Moreira Veloso*  
*Divino Batista de Souza*

**1ª edição**

1ª impressão (2009): tiragem 100 exemplares  
Edição online (2009)

**Todos os direitos reservados**

A reprodução não-autorizada desta publicação, no todo ou em parte, constitui violação dos direitos autorais (Lei no 9.610).

**Dados Internacionais de Catalogação na Publicação (CIP)**  
**Embrapa Cerrados**

---

G943 Guia de utilização do software Alceste: uma ferramenta de análise lexical aplicada à interpretação de discursos de atores na agricultura/ Evie dos Santos de Sousa... [et al.]. – Planaltina, DF : Embrapa Cerrados, 2009.  
37 p.– (Documentos / Embrapa Cerrados, ISSN 1517-5111, ISSN online 2176-5081 ; 275).

1. Software. 2. Análise lexical. 3. Agricultura. I. Evie dos Santos de Sousa. II. Série.

005.3 - CDD 21

---

© Embrapa 2009

# **Autores**

**Evie dos Santos de Sousa**

Engenheira Agrônoma, M.Sc.  
Analista da Embrapa Cerrados  
evie@cpac.embrapa.br

**Márcio Alexandre dos Santos Rodrigues**

Analista de Sistema, Bacharel em Sistemas de Informação  
Bolsista da Embrapa Cerrados  
marcio.pretin@gmail.com

**Francisco Eduardo de Castro Rocha**

Engenheiro Agrícola, Psicólogo, D.Sc.  
Pesquisador da Embrapa Cerrados  
rocha@cpac.embrapa.br

**Cíntia Ribeiro Martins**

Psicóloga, M.Sc.  
Professora da Faculdade Nobre de Feira de Santana  
Feira de Santana, BA  
crlmartinspsi@gmail.com.

# Agradecimentos

Agradecemos ao Conselho Nacional de Desenvolvimento Científico e Tecnológico pelo apoio para importação do software Alceste; à Embrapa Cerrados pela contratação do bolsista de Sistemas de Informação que auxiliou no entendimento deste software; ao Setor de Patrimônio e Materiais desta Unidade pelo desenvolvimento do processo de compra do software e à Área de Comunicação e Negócios desta Unidade pela confiança dada ao Núcleo de Pesquisa Transversal para o desenvolvimento dos trabalhos de Avaliação de Programas no contexto da Transferência de Tecnologia.

# Apresentação

Há uma mudança em curso na Embrapa que precisa ser melhor diagnosticada, compreendida e avaliada. Ela é típica de períodos de transição, em que novos valores e atitudes desmontam lenta e progressivamente o que havia antes. Dessa vez, há um agravante: em nenhum outro momento na história da Empresa, a mudança foi tão urgente e a sociedade tão demandante desta mudança.

A mudança exige que a Embrapa desenvolva tecnologias que protejam o meio ambiente, a qualidade de vida das pessoas, solucionem os problemas de alimentação e saúde, incluam produtores familiares no processo produtivo, gerem fontes de matéria-prima e energia e conseqüentemente riquezas para o País.

Nesse contexto, é necessário melhorar as formas de avaliação dos resultados gerados pela empresa. Há necessidade de entender que fatores condicionam a adoção de tecnologias, identificar os indicadores de impacto das tecnologias transferidas para a sociedade e avaliar as intervenções técnicas promovidas pela empresa no processo de transferência de tecnologia. Este guia apresenta um instrumento que poderá contribuir para essas avaliações, o software Alceste.

O guia está dividido em capítulos didáticos de forma sintetizada: o que é o software, seu histórico e aplicação, a configuração do computador, o processo de instalação e utilização, a preparação do *corpus*, as possibilidades de análise e seus produtos, as etapas de análise realizadas e os conceitos básicos que orientarão a sua leitura. Sem prejuízo do realismo dos textos, este programa serve para analisar grandes volumes de bancos de dados textuais de forma eletrônica, rápida e automática. Dessa análise, as ideias essenciais são identificadas constituindo aspectos fundamentais do discurso dos participantes da avaliação de um determinado objeto.

De posse desses aspectos, o pesquisador poderá organizar o discurso dos participantes e avaliar a eficiência, eficácia e efetividade de programas sociais, intervenções técnicas de transferência de tecnologia, identificar indicadores de impacto de tecnologias e avaliar a adoção de tecnologias geradas pela Embrapa em benefício da sociedade brasileira.

*José Robson Bezerra Sereno*  
Chefe-Geral da Embrapa Cerrados

# Sumário

Introdução.....	11
O que é o Alceste?.....	12
Histórico e Aplicação do Alceste.....	13
Configuração do Computador.....	13
Processo de Instalação do Software Alceste.....	13
Utilização do Alceste.....	17
Conceitos Básicos.....	32
Preparação do <i>Corpus</i> .....	33
Possibilidades de Análise e seus Produtos.....	34
Etapas de Análise Realizadas pelo Software.....	35
Considerações Finais.....	35
Referências.....	36
Abstract.....	37

# Guia de Utilização do Software Alceste: uma ferramenta de análise lexical aplicada à interpretação de discursos de atores na agricultura

---

*Evie dos Santos de Sousa*

*Márcio Alexandre dos Santos Rodrigues*

*Francisco Eduardo de Castro Rocha*

*Cíntia Ribeiro Martins*

## Introdução

Em razão da crescente demanda por estudos relacionados à avaliação da adoção e impacto de tecnologias e práticas conservacionistas na Embrapa, novos métodos de pesquisa, aportes teóricos e ferramentas de trabalho têm sido objeto de interesse e utilização. Conhecer a opinião dos usuários/clientes envolvidos nesse contexto torna-se base para a compreensão dos fatores determinantes do uso e das consequências relacionados às atividades dos pesquisadores envolvidos na geração de tecnologias e serviços.

Esse tipo de pesquisa pode compreender tanto estudos qualitativos quanto quantitativos e, às vezes, os dois (COZBY, 2003; MINAYO, 1996). No que se refere ao estudo qualitativo, a análise de discursos de atores sociais, por meio da técnica da análise de conteúdo, é bastante empregada e precisa, especialmente no que se refere ao levantamento de indicadores de avaliação, que nem sempre são encontrados na literatura e são relevantes para a construção de instrumentos de mensuração dos diferentes fenômenos sociais que envolvem tanto as tecnologias quanto as referidas práticas conservacionistas (MINAYO, 1999; RICHARDSON, 1999; SOUZA; GOMES, 2003; DENZIN; LINCOLN, 2006; ROCHA et al., 2008).

Estudos qualitativos e quantitativos incluem as seguintes etapas de execução: (1) entrevistas abertas, aplicadas individualmente e gravadas, com especialistas e público-alvo relacionados com o problema de pesquisa; (2) transcrição das entrevistas para o formato de texto; (3) análise das informações transcritas por meio da técnica de análise de conteúdo, que pode ser tanto manual (quando se está diante de pequenos textos) ou eletrônica (quando se está diante de grandes volumes de textos), que pode ser feita por meio do software Alceste; (4) elaboração de instrumento de mensuração do objeto de pesquisa (questões fechadas), elaborados por meio de escalas de 5 ou 10 pontos, tipo Likert; (5) análise dos dados quantitativos por meio de softwares como o Statistical Package for the Social Sciences – SPSS (ZANELLI, 2002; TRIVINOS, 2006).

As análises de conteúdos textuais podem ser feitas de forma manual com base nos significados envolvidos nos textos (semântica) ou eletrônica com base na estrutura gramatical (léxico) (NASCIMENTO; MENANDRO, 2006). Os textos podem ser obtidos de entrevistas com grupos de pessoas, especialistas, agricultores, empregados, capítulos de livros, livros, documentos, relatórios, revistas, ou seja, em qualquer base textual desde que seja de interesse de pesquisa da Embrapa. Dessa forma, a análise de conteúdo pode ser aplicada em pesquisas sobre desenvolvimento institucional, de opinião pelas áreas de comunicação e negócios, na pesquisa transversal em transferência de tecnologia que envolve avaliação da adoção e de impactos das tecnologias pela Unidade.

O fundamental para realizar esse tipo de análise é dispor de um banco de dados bem estruturado para retirar do mesmo as principais informações sobre o objeto de estudo. Para bancos de dados textuais pequenos, o recomendado é realizar a análise de conteúdo manual e para grandes volumes textuais, a análise de conteúdo eletrônica feita com o software Alceste. Esse software permite a rápida análise de dados e retirada de informações úteis à pesquisa, como, por exemplo, indicadores qualitativos e quantitativos e categorias referentes ao objeto de estudo (MARTINS, 2008).

Para exemplificar, recentemente a Embrapa Cerrados realizou um estudo sobre o uso racional da água na agricultura, considerando as opiniões de extensionistas rurais da Emater DF, que foram entrevistados a partir de um roteiro definido e todas as entrevistas foram gravadas e transcritas para compor um banco de dados textual. Para análise desse banco, foi usado o software Alceste, que forneceu, em menos de um minuto, as categorias, as unidades de recorte do texto e suas respectivas frequências.

O uso do software permitiu identificar as crenças compartilhadas pelos extensionistas acerca do uso sustentável da água na agricultura, revelando que categoria estaria associada a esse uso e os léxicos mais representativos de cada uma em termos de frequências e percentuais. Nesse estudo, o léxico "Irrig + " foi o mais representativo.

A análise de conteúdo manual desse banco de dados também foi realizada, uma vez que o volume de dados não era grande, permitindo comparar essa análise com a eletrônica. Os resultados de análise manual foram relacionados ao da eletrônica. Dessa forma, as análises se complementam conforme Nascimento e Menandro (2006).

Também foi feita uma pesquisa relacionada à gestão da Unidade que abrangeu analistas e pesquisadores da Unidade localizados fisicamente na plataforma II sobre a proposta de mudança de suas salas para a plataforma I. Esse público foi convidado a responder um questionário visando verificar suas crenças acerca dessa mudança. As respostas formalizaram um banco de dados digital que foi analisado pelo Alceste. Os resultados permitiram identificar fatores relacionados à mudança que foram esclarecidos pela Chefia da Unidade por meio de intervenções (reuniões) com o público entrevistado.

O Alceste é um software de análise de banco de dados textuais que permite a avaliação de objetos com base nas opiniões de participantes dessa avaliação (CAMARGO, 2005). Para realizar qualquer avaliação, é preciso medir e, para isso, ter instrumentos confiáveis.

Nesse contexto, o uso do Alceste como um instrumento confiável de análise de dados textuais obtidos das opiniões dos participantes de estudos promovidos pelo Núcleo de Pesquisa Transversal da Embrapa Cerrados, com o objetivo de avaliar a adoção de tecnologias, identificar os impactos (social, econômico e ambiental) das tecnologias e oferecer *feedback* para a pesquisa sobre tais avaliações, permitindo que o processo de transferência de tecnologia e de pesquisa e desenvolvimento se beneficiem dos resultados.

Dessa forma, o Alceste pode ser aplicado na Embrapa em qualquer estudo que envolva dados textuais em conjunto ou não com a análise de conteúdo manual, mas com a finalidade de identificar os principais léxicos presentes no banco de dados.

## O que é o Alceste?

O *software Alceste (Analyse lexicale par context d'un ensemble de segments de texte)* é um programa que permite realizar a análise de dados textuais ou as análises de estatística e matemática verificando a principal informação presente no texto (CAMARGO, 2005). A análise de um conjunto

de segmentos de texto está baseada na identificação das palavras (o léxico) mais significativas. Isso permite a análise de grandes volumes de dados textuais de forma automática. Esse software pode ver conjugado à análise de conteúdo clássica Bardin (2004), complementando essa análise.

O objetivo da aplicação do software, como técnica de análise de dados textuais, permite quantificar o texto para extrair estruturas que apresentem significados mais representativos e acessar a informação essencial presente no texto. Isso permite descrever, classificar, assimilar, sintetizar e identificar a organização tópica de um texto, acessando as relações existentes entre os léxicos.

## Histórico e Aplicação do Alceste

O software Alceste foi desenvolvido por Max Reinert, na França, na década de 1970, para analisar textos. Tradicionalmente, foi aplicado em estudos de Linguística e, apesar de originalmente trabalhar com a língua francesa, possui dicionários em outros idiomas, o que permite sua utilização com materiais em português. É desenvolvido e distribuído pela empresa Image na França, especializada em desenvolvimento de softwares científicos e aplicações matemáticas. No Brasil, o Alceste foi introduzido em 1998 (CAMARGO, 2005).

Atualmente, é usado nas Ciências Humanas e Sociais (linguística, psicologia, sociologia, história), para analisar textos que preservam a qualidade dos fenômenos estudados por essas Ciências). Também pode ser usado para analisar bancos de dados textuais oriundos de comunicação escrita (artigos da mídia, textos literários, documentos, relatórios, entre outros) ou transcrita (falas de entrevistas, diálogos, depoimentos, relatos, questões abertas de questionários) (CAMARGO, 2005).

O software pode ser usado na avaliação de programas sociais (CANO, 2004) e em intervenções técnicas usando as opiniões de pessoas envolvidas com o objeto a ser avaliado (ROCHA et al., 2008). Na Embrapa, o programa poderá ser aplicado na avaliação da adoção de tecnologias geradas pela empresa; de intervenções técnicas realizadas pela transferência de tecnologia e no levantamento de indicadores de avaliação de impacto/resultados com base nas opiniões das pessoas (usuários de tecnologias, pessoas que recebem as intervenções, como produtores, técnicos de assistência técnica, de empresas privadas, parceiros técnicos, entre outros).

## Configuração do Computador

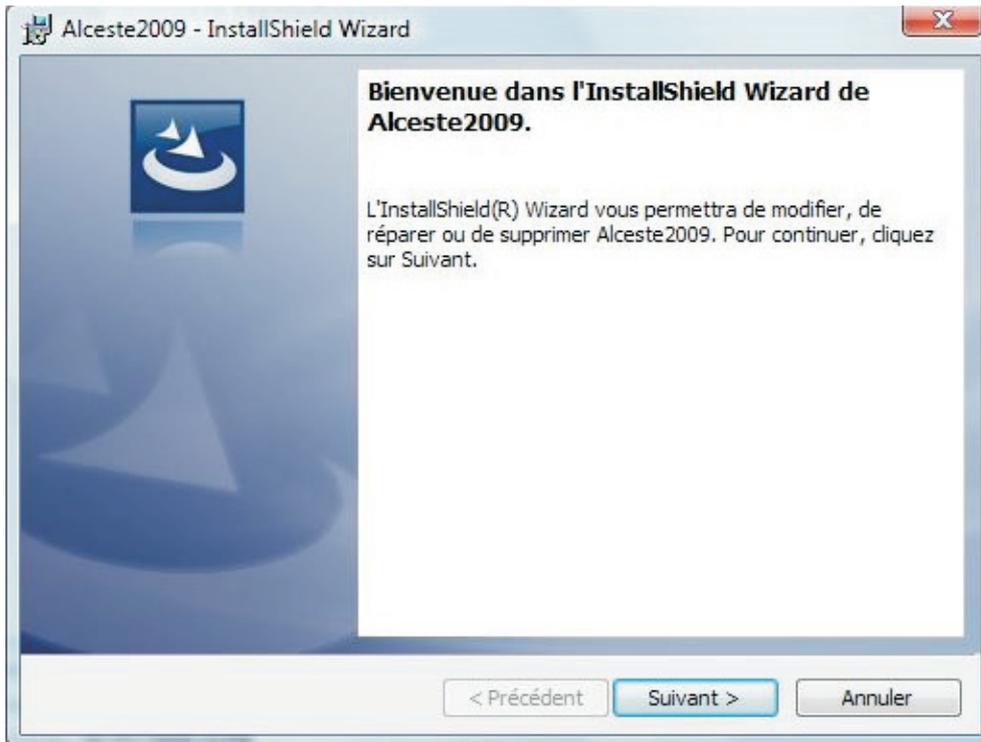
O computador deverá conter como principais características mínimas: processador Pentium IV; sistema operacional Windows XP; memória RAM de 1 Gbytes e 180 MB de disco rígido; resolução de vídeo de 1024 x 768 e sistema de cores médio (16 bits). A capacidade de análise pode ser de 1 MB, de 3 MB ou 15 MB, dependendo da versão do Alceste.

## Processo de Instalação do Software Alceste

Para a instalação do *software* Alceste, deve-se seguir as seguintes instruções:

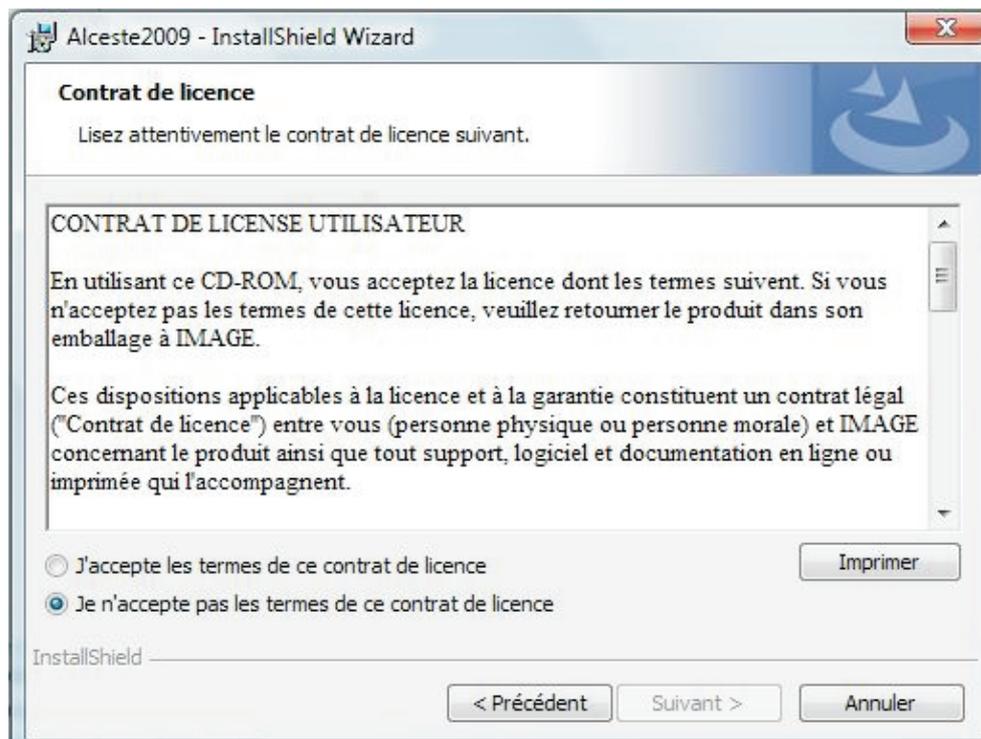
1. O Alceste pode ser instalado por meio de CD de instalação presente nas versões antigas, bem como por meio de download do *software* e *pendrive* específicos para a versão mais atualizada que é Alceste 2009 (IMAGE, 2009).

O arquivo dessa versão deve ser baixado no site: [www.image-zafar.com](http://www.image-zafar.com), conforme instruções do fabricante encaminhadas com o *pendrive* na compra do software. Após o *download*, dando-se dois cliques no arquivo "ALCESTE2009.exe" aparecerá a seguinte tela:



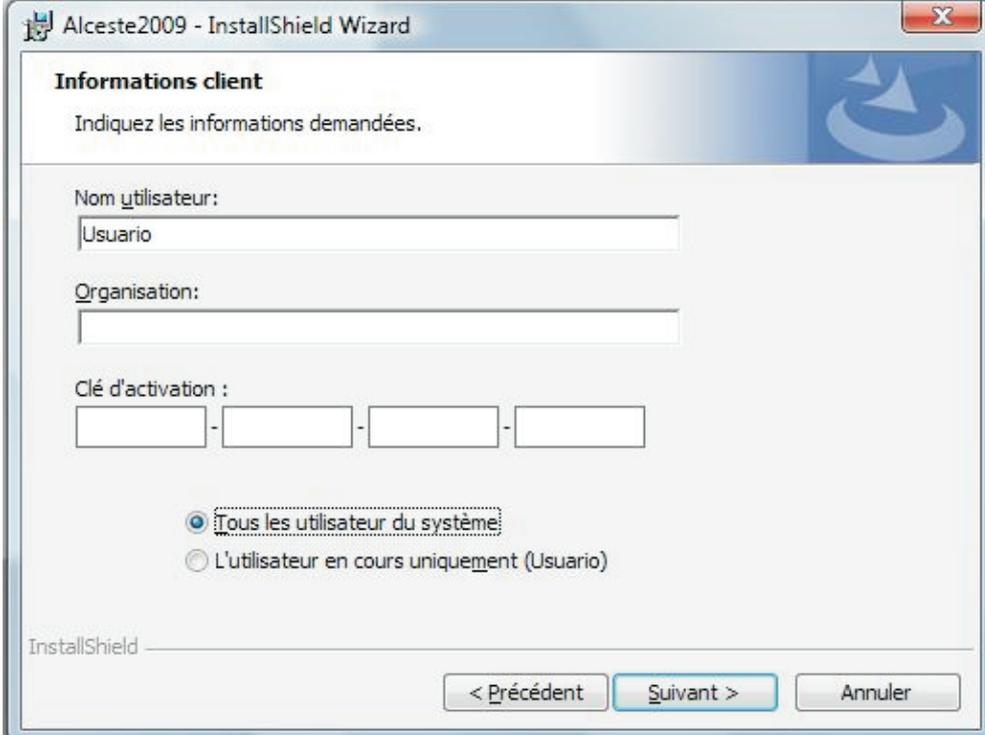
**Comentário:** Clique em "Suivant >" (Próximo), para avançar na instalação.

2. Aparecerá outra tela contendo o contrato de licença de uso do *software* que deverá ser lido atentamente, podendo ser impresso. Após sua leitura, para prosseguir com a instalação, clique em "J'accepte les termes de ce contrat de licence" (Eu aceito os termos deste acordo de licença).



**Comentário:** Clique em "Suivant >" (Próximo), para avançar na instalação.

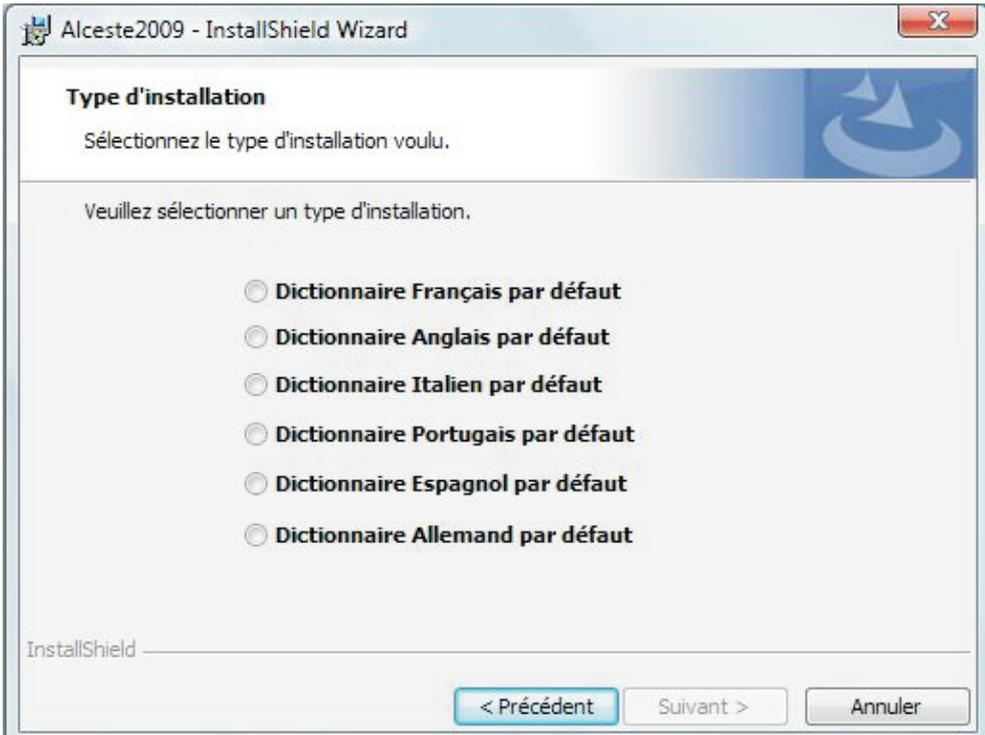
3. Colocar o nome do usuário, o da empresa a qual pertence e a chave de ativação (clé d'activation), que são quatro sequências de números e letras enviados pelo fabricante juntamente com o *pendrive* ou chave de proteção (*clé protection*). Após a inserção desses dados, deve-se inserir o *pendrive* na porta *USB*.



The screenshot shows the 'Alceste2009 - InstallShield Wizard' window. The title bar reads 'Alceste2009 - InstallShield Wizard'. The main window has a blue header with the text 'Informations client' and 'Indiquez les informations demandées.' Below this, there are three input fields: 'Nom utilisateur:' with the text 'Usuario', 'Organisation:', and 'Clé d'activation:' which is a four-part field with dashes. There are two radio button options: 'Tous les utilisateur du système' (which is selected) and 'L'utilisateur en cours uniquement (Usuario)'. At the bottom, there are three buttons: '< Précédent', 'Suivant >', and 'Annuler'.

**Comentário:** Clique em "Suivant >" (Próximo), para avançar na instalação.

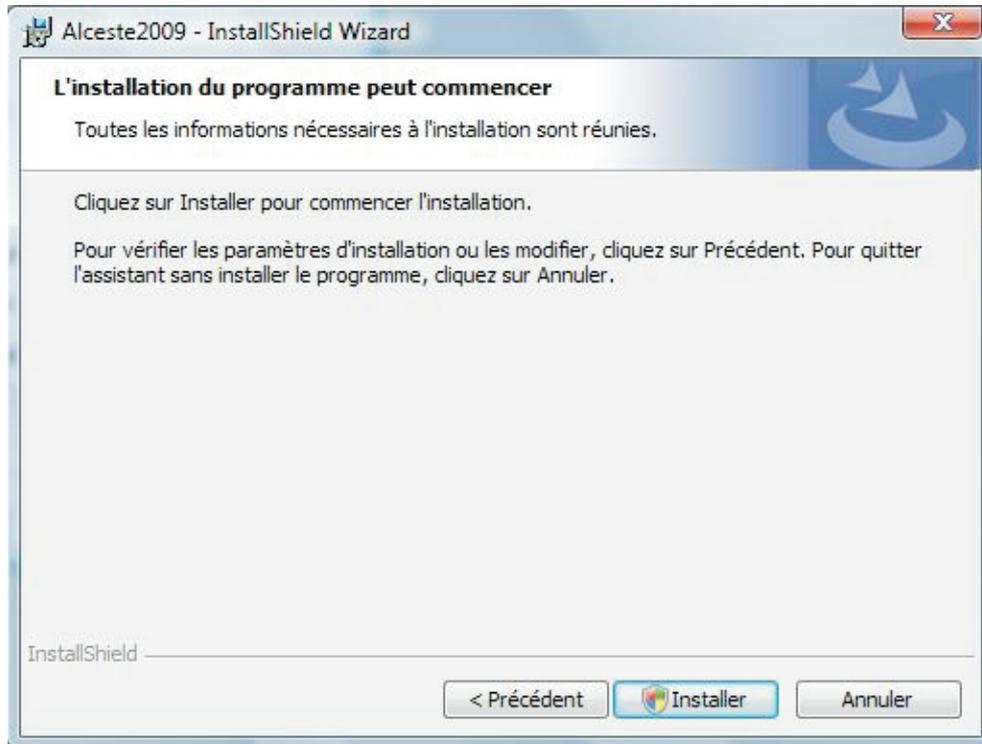
4. Em seguida, deve-se escolher o dicionário que melhor atenda a linguagem do arquivo a ser analisado. Para um arquivo em português, a opção marcada deverá ser "*Dictionnaire Portugais par défaut*".



The screenshot shows the 'Alceste2009 - InstallShield Wizard' window. The title bar reads 'Alceste2009 - InstallShield Wizard'. The main window has a blue header with the text 'Type d'installation' and 'Sélectionnez le type d'installation voulu.' Below this, there is a text prompt 'Veuillez sélectionner un type d'installation.' and a list of six radio button options: 'Dictionnaire Français par défaut', 'Dictionnaire Anglais par défaut', 'Dictionnaire Italien par défaut', 'Dictionnaire Portugais par défaut', 'Dictionnaire Espagnol par défaut', and 'Dictionnaire Allemand par défaut'. At the bottom, there are three buttons: '< Précédent', 'Suivant >', and 'Annuler'.

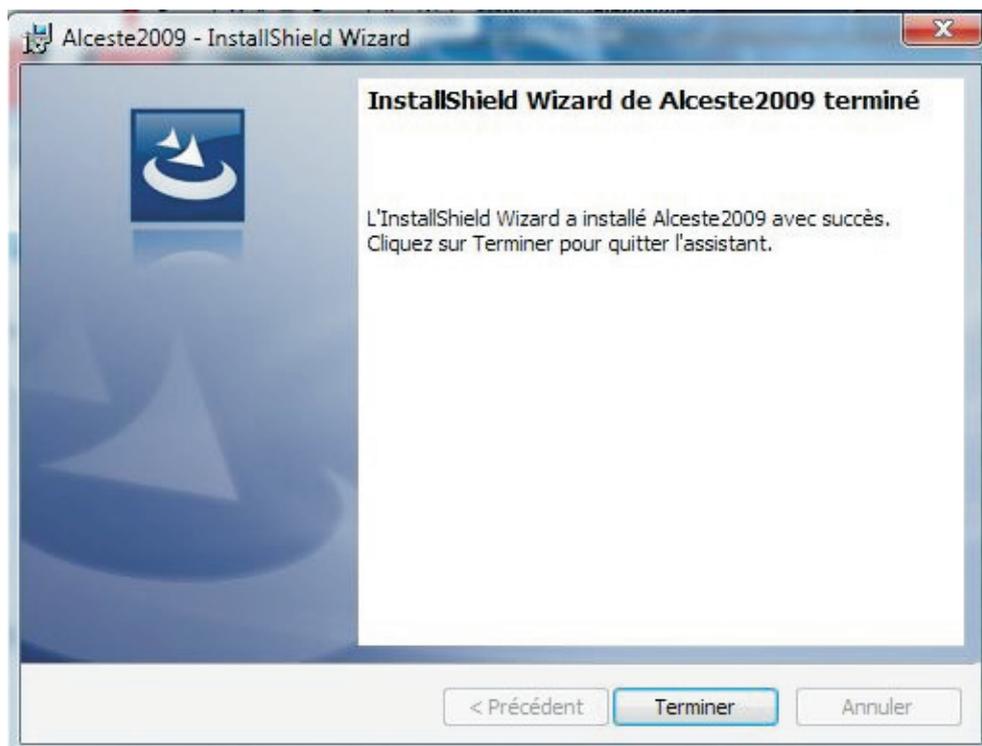
**Comentário:** Clique em "Suivant >" (Próximo), para avançar na instalação.

5. Aparecerá, em seguida, a tela informando que a instalação do programa será iniciada. Clique no Botão "Installer" para continuar a instalação.



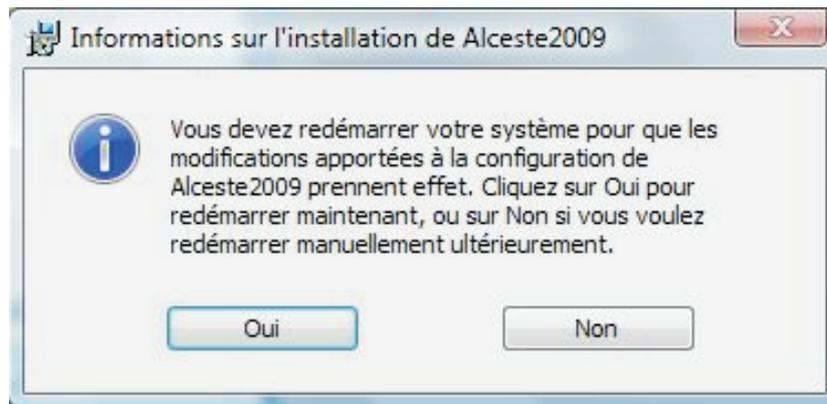
**Comentário:** o tempo de instalação dependerá da configuração do computador do usuário.

6. Uma vez concluída a instalação, clique em "Terminer".



**Comentário:** clique em "Terminer" para completar a instalação.

7. Reiniciar o computador clicando em "Oui", conforme tela abaixo.



**Comentário:** ao reiniciar o computador, são mantidas as configurações do programa.

## Utilização do Alceste

1. Para inicializar o uso do *software*, deve-se dar dois cliques no seguinte ícone que aparecerá no *desktop* do computador:



**Comentário:** após reiniciar o computador, esse ícone automaticamente estará na área de trabalho.

2. Aparecerá, em alguns segundos, a tela principal do Alceste:



**Comentário:** essa tela inicial é típica da versão 2009.

3. No menu da tela inicial, têm-se os seguintes itens:

3.1. O primeiro item do menu é o *Fichier* (arquivo), que contém 2 subitens de nome *Préférences* (Preferências) e *Quitter* (Sair).



3.2. O segundo item *Traitement* (tratamento) tem os subitens *Nouveau* (Novo) e *Ouvrir* (Abrir), conforme a figura abaixo.

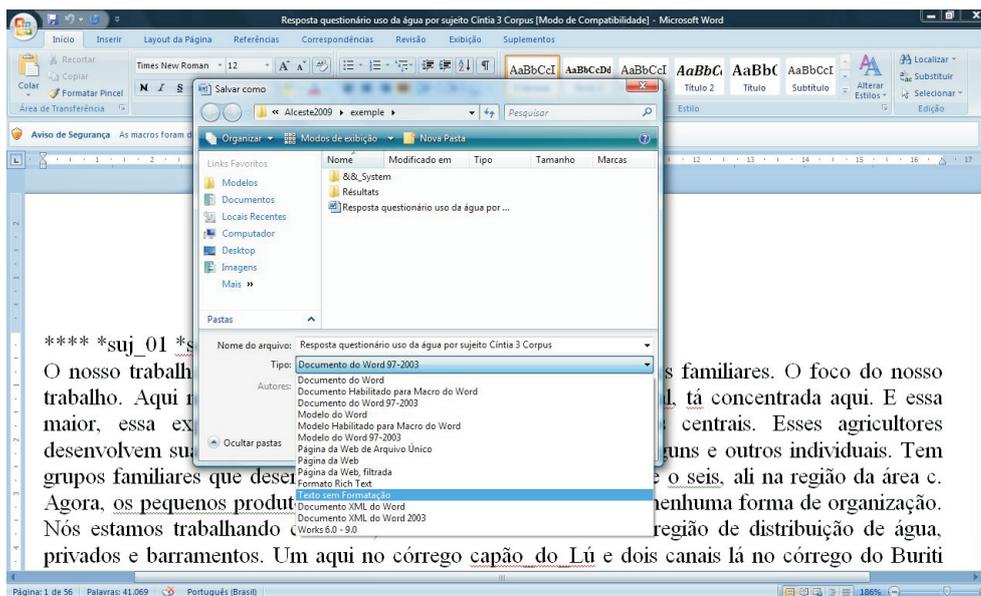


3.3. O terceiro e último item do menu "?" (ajuda) tem os subitens "A propos d'ALCESTE" (sobre o Alceste) e *Remerciements* (Agradecimentos).

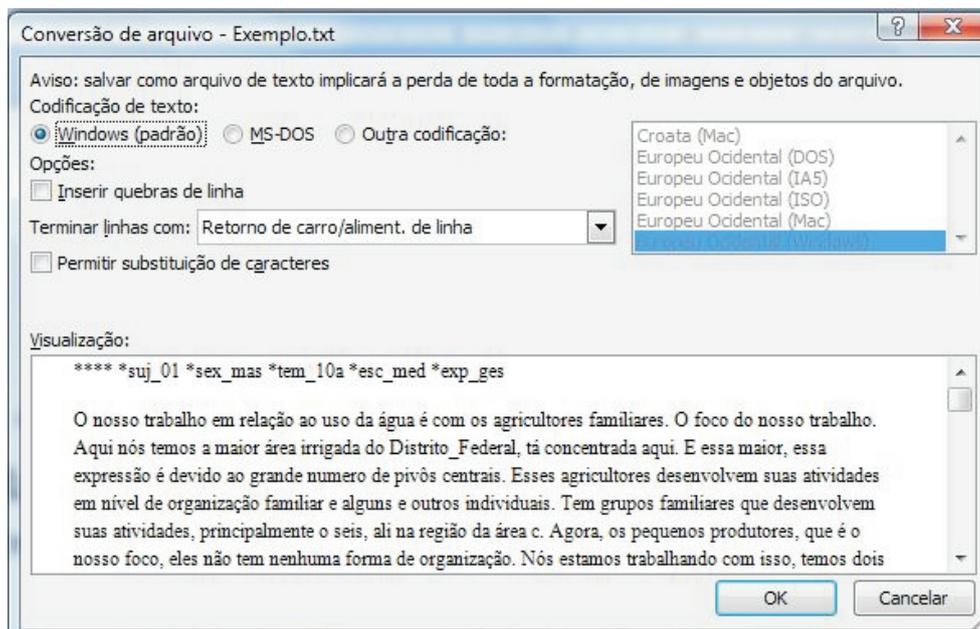


Para começarmos a utilizar o programa, serão feitas as seguintes operações:

4. O primeiro passo é preparar o *corpus* (nome comum dado ao arquivo com os dados a serem analisados) na extensão ".doc". Depois, o arquivo deverá ser transformado para o formato ou extensão ".txt" da seguinte forma:
  - 4.1. Usando o *Microsoft Word*, salve o arquivo que contém os dados textuais a serem analisados usando o "Salvar como...". Após a abertura da janela, deve-se abrir a aba chamada "Tipo:" e escolher o item "Texto sem formatação". Depois, procure a aba "Nome do arquivo:" e mude o nome do seu arquivo para um nome pequeno (use no máximo 7 caracteres), pois o programa não aceita nomes extensos. Neste tutorial, usa-se o nome "Exemplo".



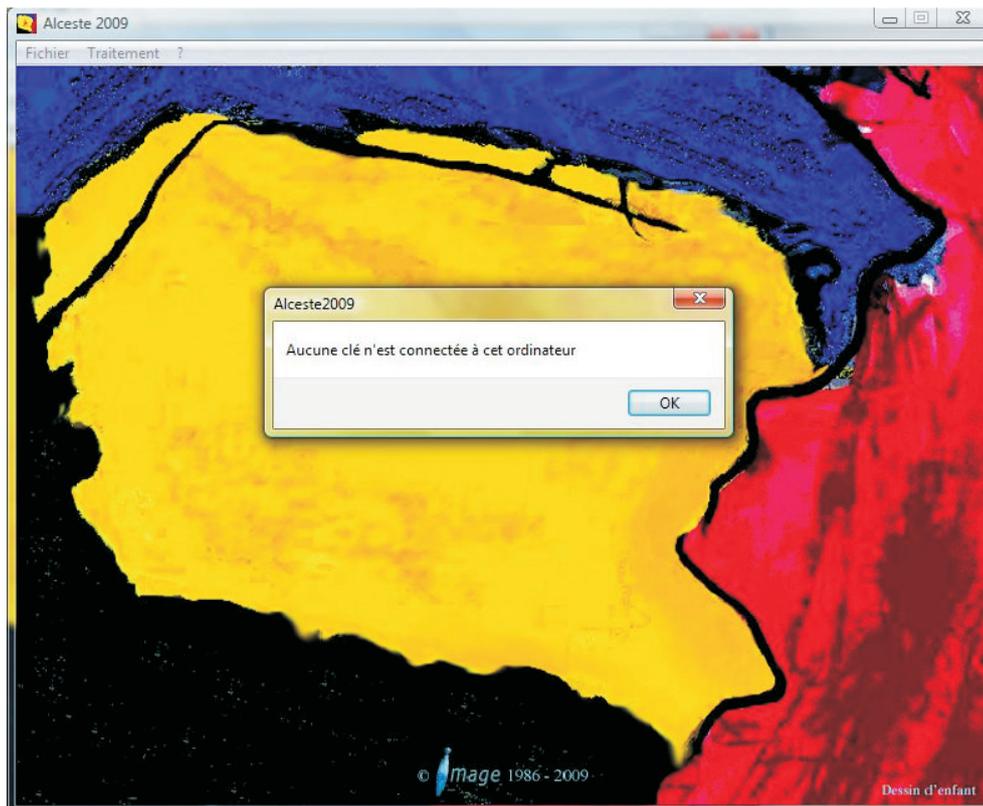
- 4.2. Aparecerá automaticamente a seguinte tela com as opções. Deve-se marcar a opção “Inserir quebras de linha”. Caso não seja marcada essa opção, o Alceste não aceitará o arquivo, informando erro nas linhas porque o *software* considera somente os 1.500 caracteres em cada linha do texto.



Após as alterações, o arquivo estará pronto para ser executado no *software*.

5. Com o *software* Alceste aberto, abra o item “*Traitement*” (Tratamento) e clique no subitem “*Nouveau*” (Novo), mostrados no subitem 3.2 deste guia.

- 5.1. Pode ocorrer a seguinte falha “*Aucune clé n’est connectée cet ordinateur*” (A chave não esta ligada a este computador),



**Comentário:** há necessidade de reiniciar a instalação do *software*.

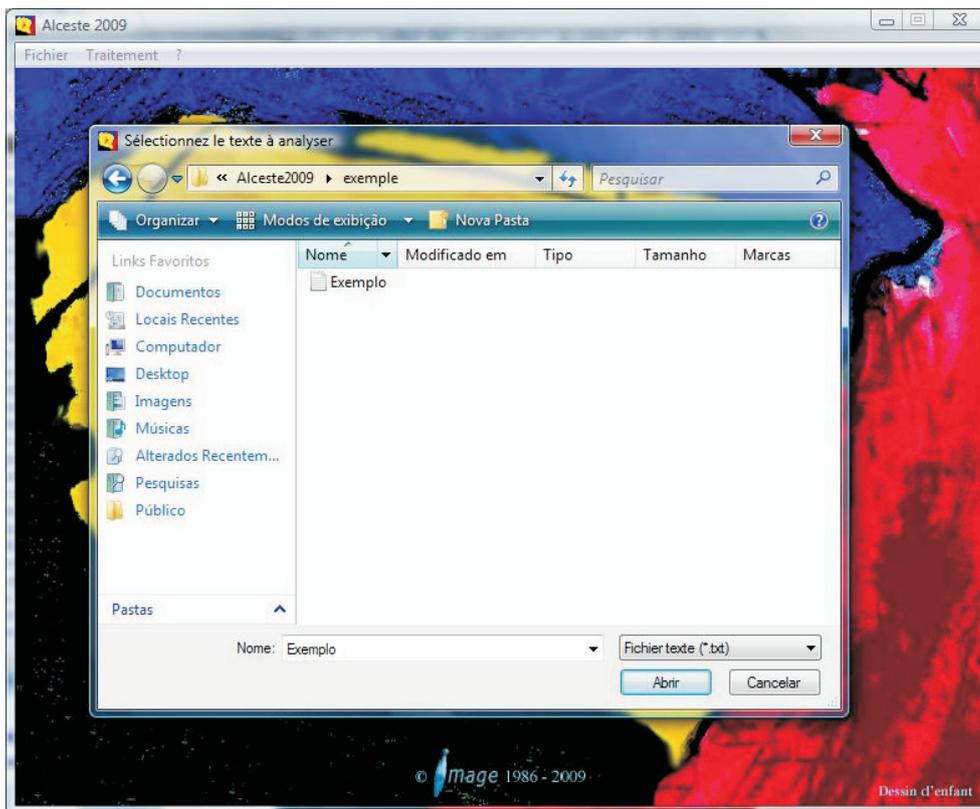
Deve-se refazer a instalação do *software* novamente, usando a chave USB que acompanha o CD de instalação. Essa chave será usada inclusive ao realizar análises.

5.2. Caso não ocorra o erro citado anteriormente, aparecerá a seguinte tela:

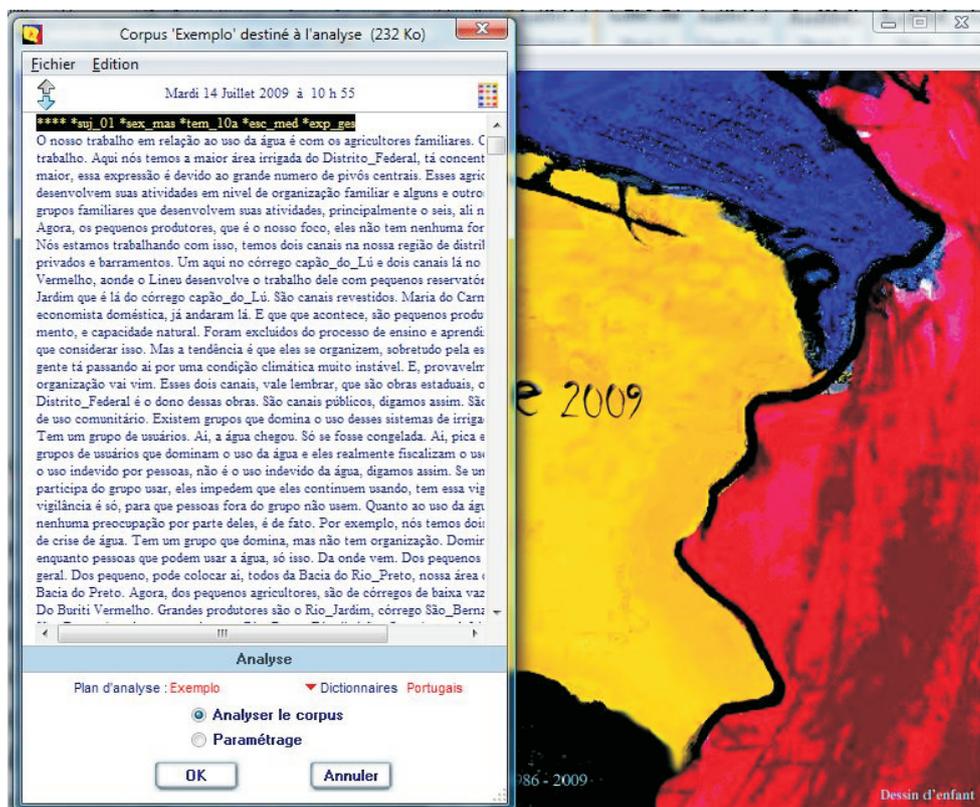


**Comentário:** essa tela confirma a correta instalação do *software*.

6. Nesta tela, escolhe-se em *Traitement* (tratamento) e *Nouveau* (Novo) o arquivo para fazer a análise, no caso, o arquivo “Exemplo.txt”, clicando no botão “Abrir”.

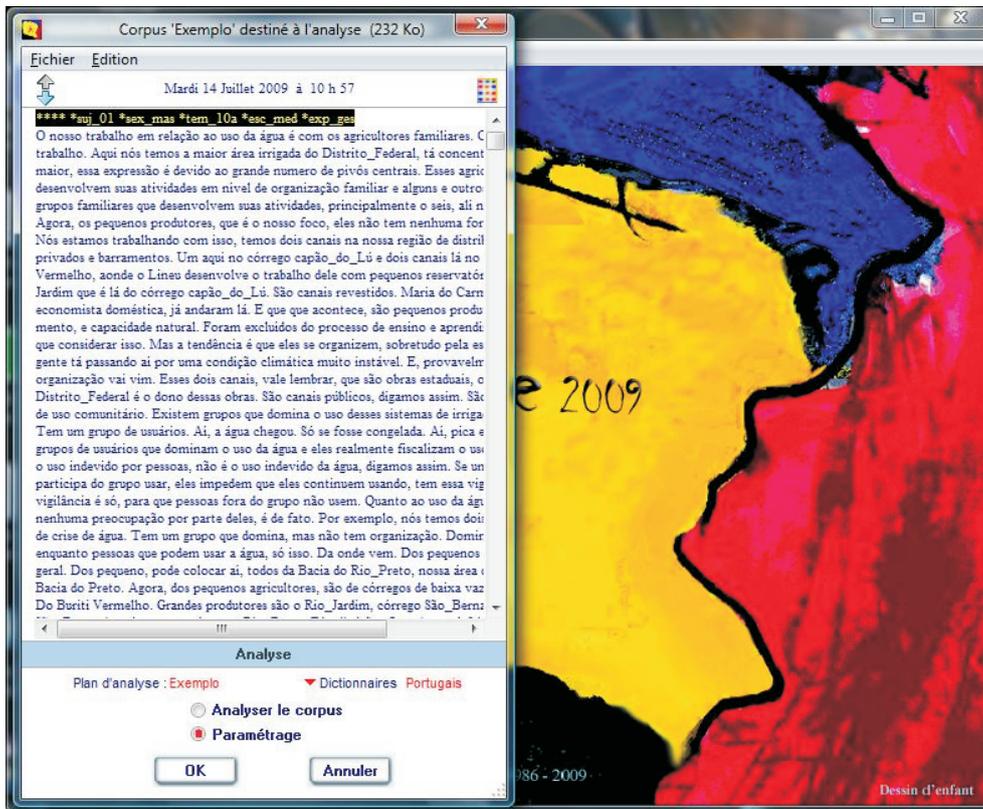


7. Após o clique, abrirá uma tela semelhante a anterior, em que aparecerá o *corpus* (nome dado pelo *software* ao arquivo em “.txt”).

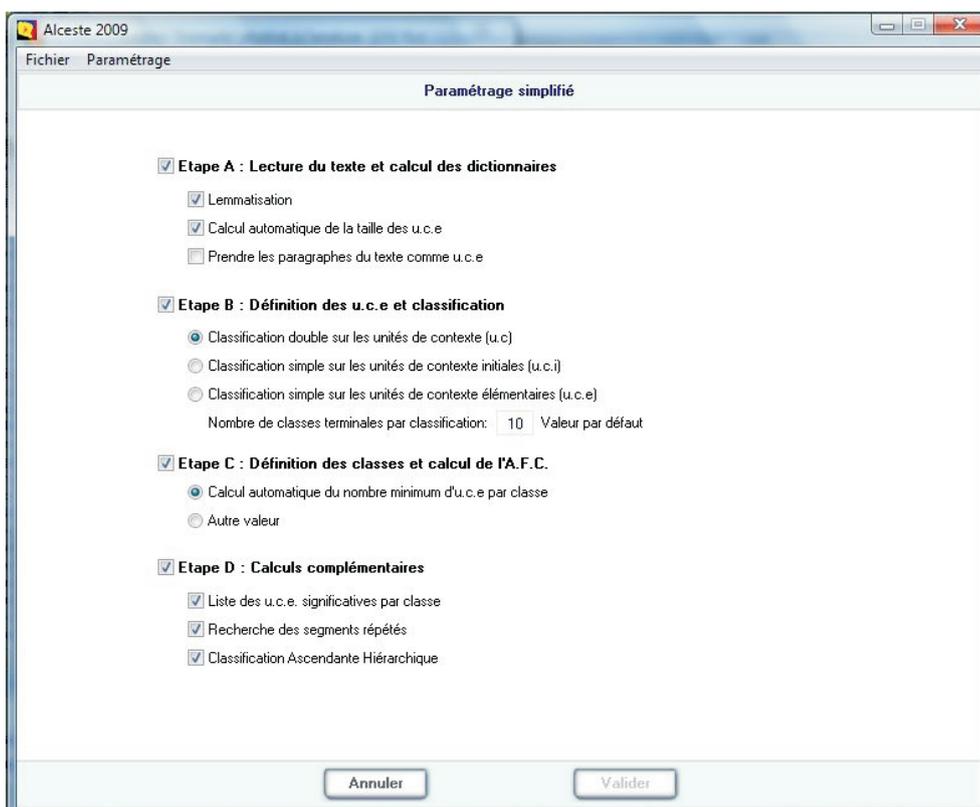


Comentário: essa tela indica que o *software* identificou o *corpus*.

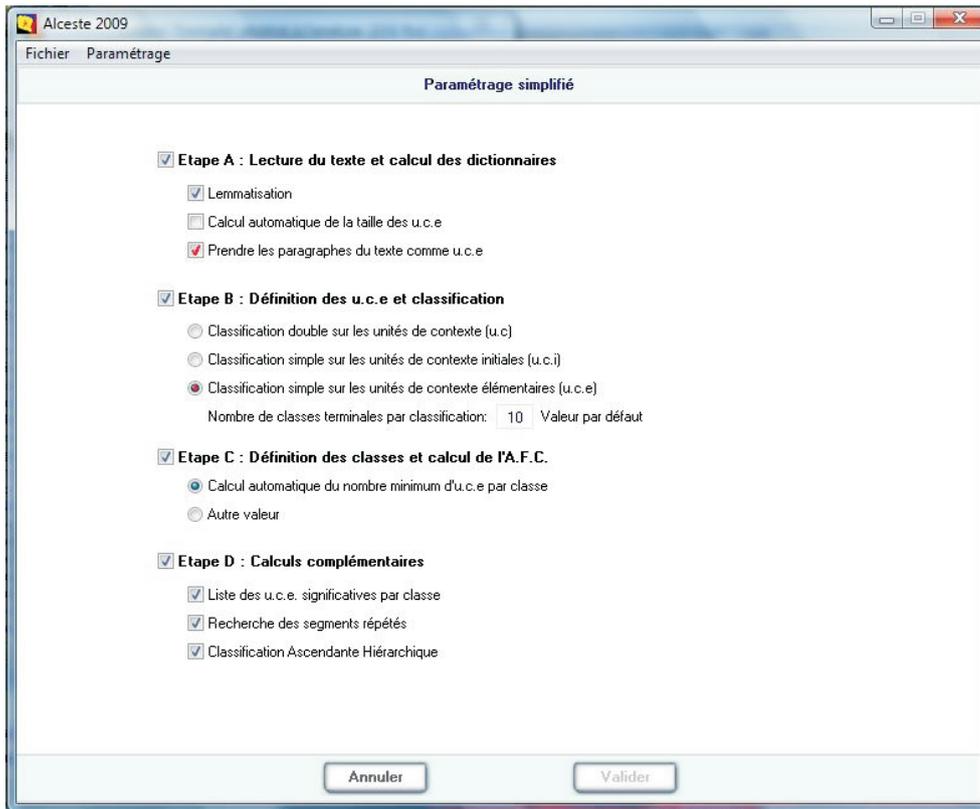
Caso sejam necessárias correções ortográficas, pode-se fazê-las nessa tela. Feitas as correções ou se não houver necessidade de correções, deve-se clicar no item “*Paramétrage*” (Parâmetros), depois clique no botão “OK” para prosseguir. Como demonstra-se a seguir:



8. Aparecerá a tela de *Paramétrage simplifié* (Parâmetros simples), para que sejam escolhidos os parâmetros de análise dos dados.



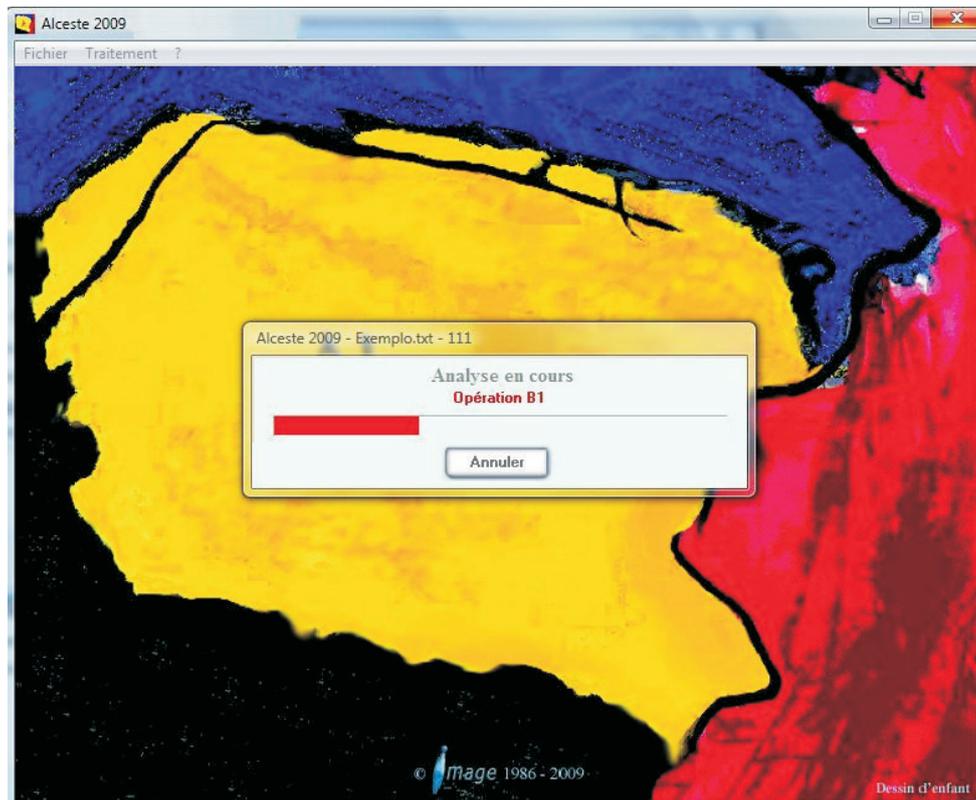
Deve-se marcar a terceira opção da Etapa A e a última opção da Etapa B, conforme a tela a seguir.



**Comentário:** o *software* começará a fazer a análise lexical do texto no arquivo (*corpus*).

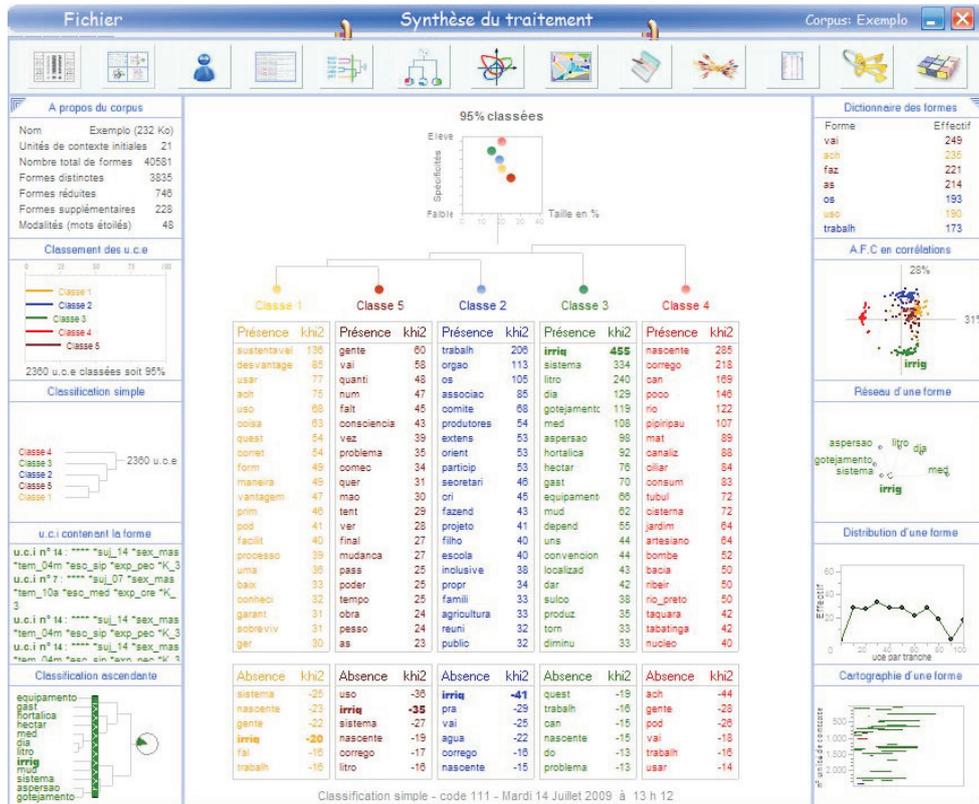
Após mudar as opções deve-se clicar no botão *Valider* para confirmar os parâmetros escolhidos.

Automaticamente, o *software* inicia a análise dos dados.



**Comentário:** as fases da análise realizada aparecem na tela.

Ao término da análise, aparecerá automaticamente a tela seguinte, com os resultados gerais da análise.

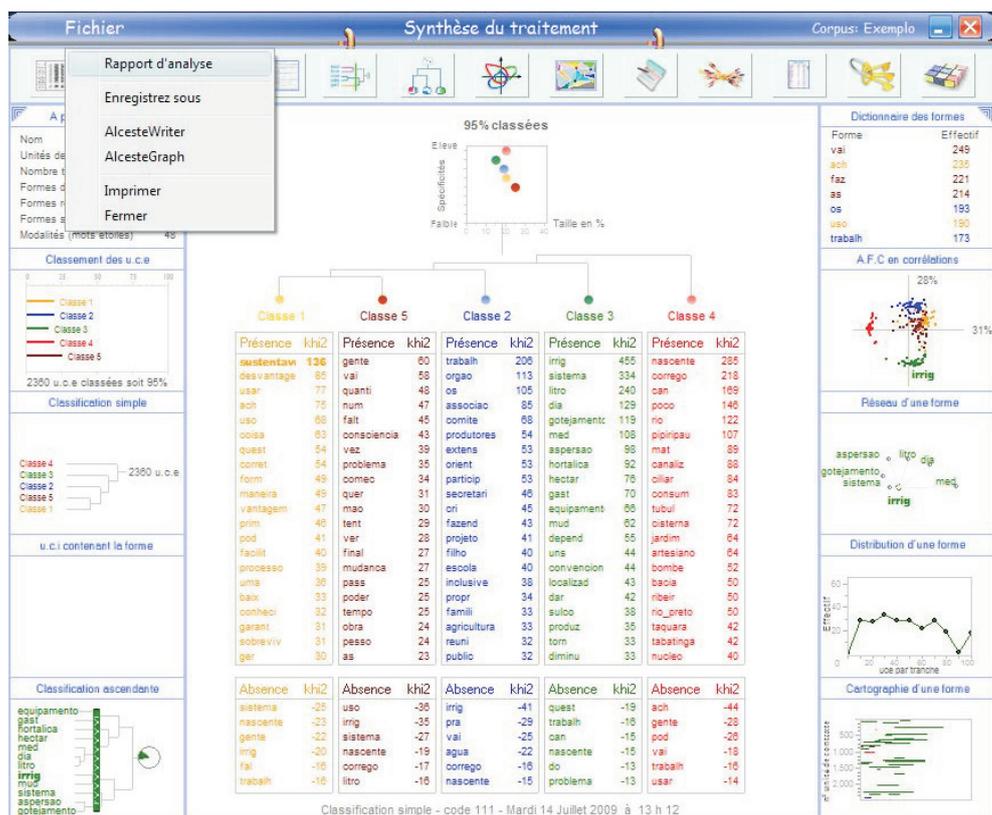


**Comentário:** a tela apresenta o conjunto sintetizado de resultados das análises.

Ao clicar sobre um exemplo de resultado, como, por exemplo, "A propous Du corpus", abre-se uma nova janela mostrando em detalhes os resultados obtidos da análise do *corpus*.

### 9. Definições do menu de análise.

Na sequência do menu dessa tela geral, seguem-se os seguintes itens: Arquivo, Análise Relatório, Salvar como, Alceste escritor, Alceste Gráfico, Imprimir, Fechar.



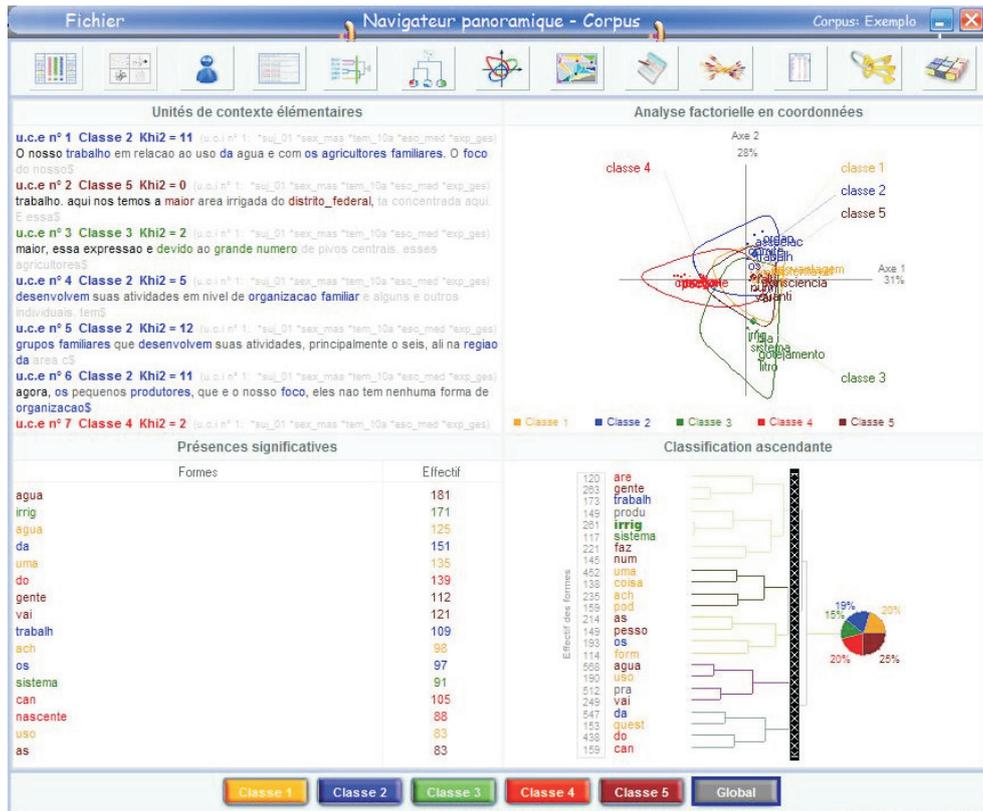
**Comentário:** a tela geral tem menu interativo.

O item relatório da análise é a saída de um relatório contendo todas as análises. Também há opções de arquivamento, salvamento e impressão.

Os documentos são lidos apenas pelo *software*, não são exportáveis.

10. Seguindo-se as funcionalidades do *software* tem-se as seguintes janelas:

10.1. *Navigateur panoramique* – Navegação Panorâmica.



Comentário: essa tela permite ver os resultados gerais e por classe identificada em separado.

10.2. *Présences* – Presenças: formas lexicais mais significativas.

The screenshot shows the 'Présences' window in Alceste, displaying a detailed table of significant forms. The table has five columns: 'Présences significatives', 'Khi2', 'Effectif de la forme', 'Pourcentage dans la classe', and 'Catégorie grammaticale'.

Présences significatives	Khi2	Effectif de la forme	Pourcentage dans la classe	Catégorie grammaticale
irrig	455	171	63%	Formes non reconnues
sistema	334	91	78%	Formes non reconnues
nascente	285	88	90%	Formes non reconnues
litro	240	57	98%	Formes non reconnues
corrego	218	67	94%	Formes non reconnues
trabalh	206	109	64%	Formes non reconnues
can	169	105	66%	Formes non reconnues
poco	146	56	87%	Formes non reconnues
mil	143	27	96%	Numéraux
por	138	103	40%	Prépositions simples et locutions prépo
sustentavel	136	45	86%	Formes non reconnues
dia	129	44	70%	Formes non reconnues
rio	122	45	83%	Formes non reconnues
gotejamento	119	27	87%	Formes non reconnues
orgao	113	30	100%	Formes non reconnues
med	108	25	86%	Formes non reconnues
pipiripau	107	31	97%	Formes non reconnues
os	105	97	50%	Formes non reconnues
aspersao	98	22	91%	Formes non reconnues
hortalica	92	24	84%	Formes non reconnues
mat	89	27	90%	Formes non reconnues
canaliz	88	22	100%	Formes non reconnues
associaç	85	26	92%	Formes non reconnues
desvantagem	85	33	87%	Formes non reconnues
ciliar	84	22	100%	Formes non reconnues
consum	83	41	71%	Formes non reconnues
usar	77	47	60%	Formes non reconnues
hectar	76	18	94%	Formes non reconnues
ach	75	98	43%	Formes non reconnues
EMATER	74	30	71%	Mots en majuscules
cisterna	72	26	100%	Formes non reconnues
tubul	72	21	100%	Formes non reconnues
gast	70	19	84%	Formes non reconnues

At the bottom, there are buttons for 'Classe 1' through 'Classe 5' and a 'Global' button.

Comentário: a tela permite ver as formas lexicais mais representativas de modo geral ou por classes identificadas em separado.

10.3. *Détail des unités de contexte* – Unidades de contexto em detalhes.

**Unités de contexte élémentaires du corpus**

u.c.e n° 1 Classe 2 KHI2 = 11 (u.c.i n° 1: "su\_01" "sex\_mas" "tem\_10a" "esc\_med" "exp\_ges")  
 O nosso trabalho em relacao ao uso da agua e com os agricultores familiares. O foco do nosso\$  
 u.c.e n° 2 Classe 5 KHI2 = 0 (u.c.i n° 1: "su\_01" "sex\_mas" "tem\_10a" "esc\_med" "exp\_ges")  
 trabalho, aqui nos temos a maior area irrigada do distrito\_federal, ta concentrada aqui. E essa\$  
 u.c.e n° 3 Classe 3 KHI2 = 2 (u.c.i n° 1: "su\_01" "sex\_mas" "tem\_10a" "esc\_med" "exp\_ges")  
 maior, essa expressao e devido ao grande numero de pivos centrais. esses agricultores\$  
 u.c.e n° 4 Classe 2 KHI2 = 5 (u.c.i n° 1: "su\_01" "sex\_mas" "tem\_10a" "esc\_med" "exp\_ges")  
 desenvolvem suas atividades em nivel de organizacao familiar e alguns e outros individuais. tem\$  
 u.c.e n° 5 Classe 2 KHI2 = 12 (u.c.i n° 1: "su\_01" "sex\_mas" "tem\_10a" "esc\_med" "exp\_ges")  
 grupos familiares que desenvolvem suas atividades, principalmente o seis, ali na regio da area c\$  
 u.c.e n° 6 Classe 2 KHI2 = 11 (u.c.i n° 1: "su\_01" "sex\_mas" "tem\_10a" "esc\_med" "exp\_ges")  
 agora, os pequenos produtores, que e o nosso foco, eles nao tem nenhuma forma de organizacao\$  
 u.c.e n° 7 Classe 4 KHI2 = 2 (u.c.i n° 1: "su\_01" "sex\_mas" "tem\_10a" "esc\_med" "exp\_ges")  
 nos estamos trabalhando com isso, temos dois canais na nossa regio de distribuicao de agua\$  
 u.c.e n° 8 Classe 4 KHI2 = 22 (u.c.i n° 1: "su\_01" "sex\_mas" "tem\_10a" "esc\_med" "exp\_ges")  
 privados e barramentos. um aqui no corrego capao\_do\_lu e dois canais la no corrego do buritis\$  
 u.c.e n° 9 Classe 4 KHI2 = 2 (u.c.i n° 1: "su\_01" "sex\_mas" "tem\_10a" "esc\_med" "exp\_ges")  
 vermelho, aonde o lineu desenvolve o trabalho dele com pequenos reservatorios. E um aqui no\$  
 u.c.e n° 10 Classe 4 KHI2 = 27 (u.c.i n° 1: "su\_01" "sex\_mas" "tem\_10a" "esc\_med" "exp\_ges")  
 jardim que e la do corrego capao\_do\_lu. sao canais revestidos. maria do carmo e nossa\$  
 u.c.e n° 11 Classe 0 KHI2 = 0 (u.c.i n° 1: "su\_01" "sex\_mas" "tem\_10a" "esc\_med" "exp\_ges")  
 economista domestica, ja andaram la. E que que acontece, sao pequenos produtores de dificil\$  
 u.c.e n° 12 Classe 1 KHI2 = 6 (u.c.i n° 1: "su\_01" "sex\_mas" "tem\_10a" "esc\_med" "exp\_ges")  
 mento, e capacidade natural foram excluidos do processo de ensino e aprendizagem. voce tem\$  
 u.c.e n° 13 Classe 5 KHI2 = 1 (u.c.i n° 1: "su\_01" "sex\_mas" "tem\_10a" "esc\_med" "exp\_ges")  
 que considerar isso, mas a tendencia e que eles se organizem, sobretudo pela escassez da agua. AS  
 u.c.e n° 14 Classe 5 KHI2 = 4 (u.c.i n° 1: "su\_01" "sex\_mas" "tem\_10a" "esc\_med" "exp\_ges")  
 gente ta passando ai por uma condicao climatica muito instavel. E, provavelmente essa\$  
 u.c.e n° 15 Classe 5 KHI2 = 0 (u.c.i n° 1: "su\_01" "sex\_mas" "tem\_10a" "esc\_med" "exp\_ges")  
 organizacao vai vim. esses dois canais, vale lembrar, que sao obras estaduais, o governo do\$  
 u.c.e n° 16 Classe 5 KHI2 = 4 (u.c.i n° 1: "su\_01" "sex\_mas" "tem\_10a" "esc\_med" "exp\_ges")  
 distrito\_federal e o dono dessas obras. sao canais publicos, digamos assim, sao canais publicos\$  
 u.c.e n° 17 Classe 4 KHI2 = 0 (u.c.i n° 1: "su\_01" "sex\_mas" "tem\_10a" "esc\_med" "exp\_ges")  
 de uso comunitario. existem grupos que domina o uso desses sistemas de irrigacao, desses canais\$  
 u.c.e n° 18 Classe 2 KHI2 = 2 (u.c.i n° 1: "su\_01" "sex\_mas" "tem\_10a" "esc\_med" "exp\_ges")  
 tem um grupo de usuarios, ai, a agua chegou. so se fosse congelada ai, pica elas assim. tem\$  
 u.c.e n° 19 Classe 2 KHI2 = 0 (u.c.i n° 1: "su\_01" "sex\_mas" "tem\_10a" "esc\_med" "exp\_ges")  
 grupos de usuarios que dominam o uso da agua e eles realmente facilitam o uso indevido. mas\$  
 u.c.e n° 20 Classe 5 KHI2 = 1 (u.c.i n° 1: "su\_01" "sex\_mas" "tem\_10a" "esc\_med" "exp\_ges")  
 o uso indevido por pessoas, nao e o uso indevido da agua. digamos assim. se uma pessoa que nao\$

**Comentário:** a tela fornece as unidades de contexto elementares gerais, por classe em separado e eliminadas. Separando aquelas com valor qui-quadrado maior, têm-se as mais representativas.

#### 10.4. Classification ascendante du corpus – Classificação ascendente do corpus.

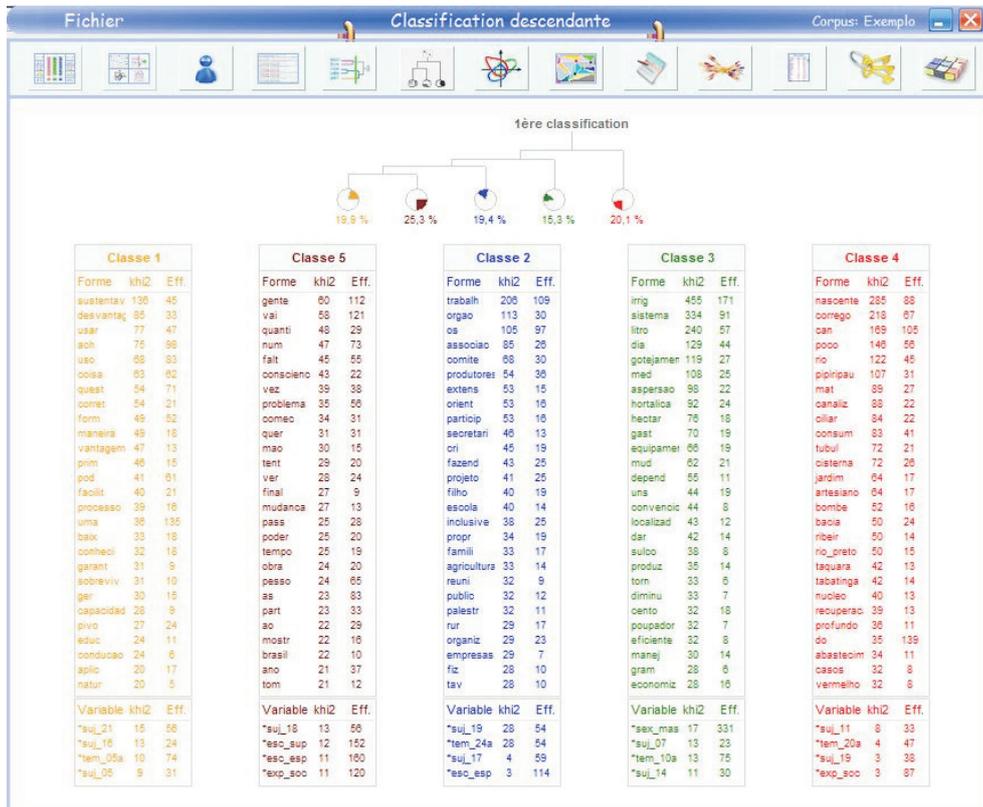
**Classification ascendante du corpus**

Effect = 120 are  
 Effect = 203 gente  
 Effect = 173 trabalh  
 Effect = 149 produ  
 Effect = 201 irrig  
 Effect = 117 sistema  
 Effect = 221 faz  
 Effect = 140 num  
 Effect = 402 uma  
 Effect = 130 coisa  
 Effect = 235 ach  
 Effect = 159 pod  
 Effect = 214 as  
 Effect = 149 pesso  
 Effect = 193 os  
 Effect = 114 form  
 Effect = 568 agua  
 Effect = 130 uso  
 Effect = 512 pra  
 Effect = 547 vai  
 Effect = 249 da  
 Effect = 100 quest  
 Effect = 430 do  
 Effect = 159 can

19% 20% 15% 20% 26%

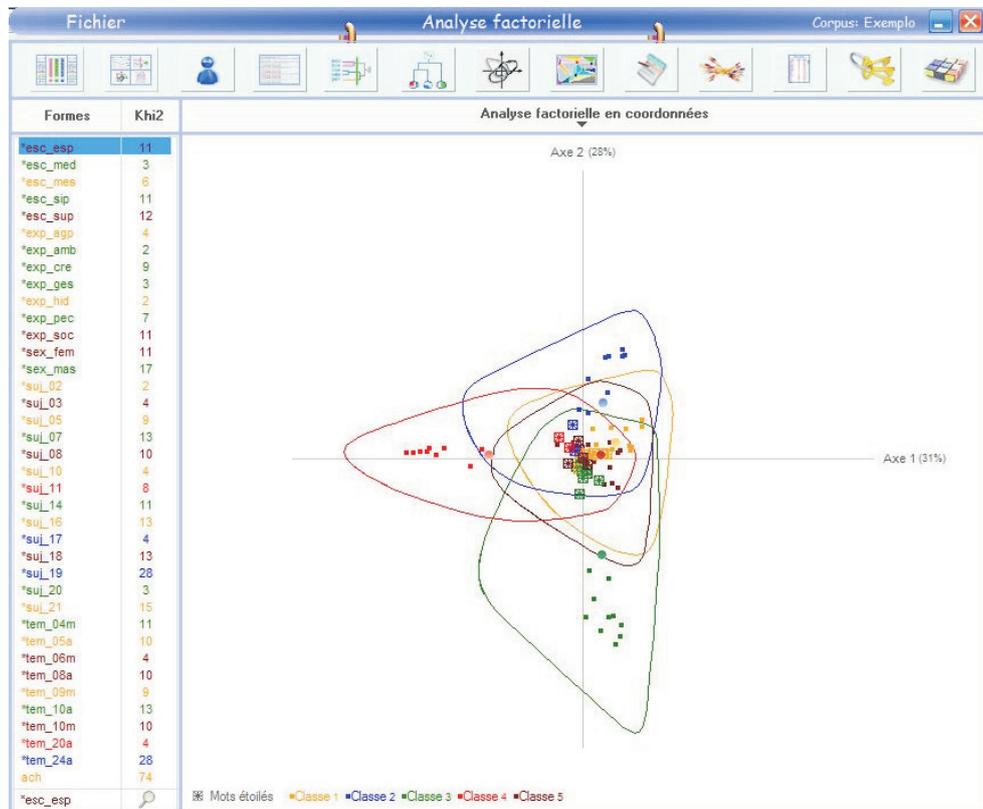
**Comentário:** resultado da classificação ascendente entre léxicos do *corpus*. Nota-se que "irrig" está próximo de "sistema" e ambos próximos de "produ".

#### 10.5. Classification descendante – Classificação descendente.



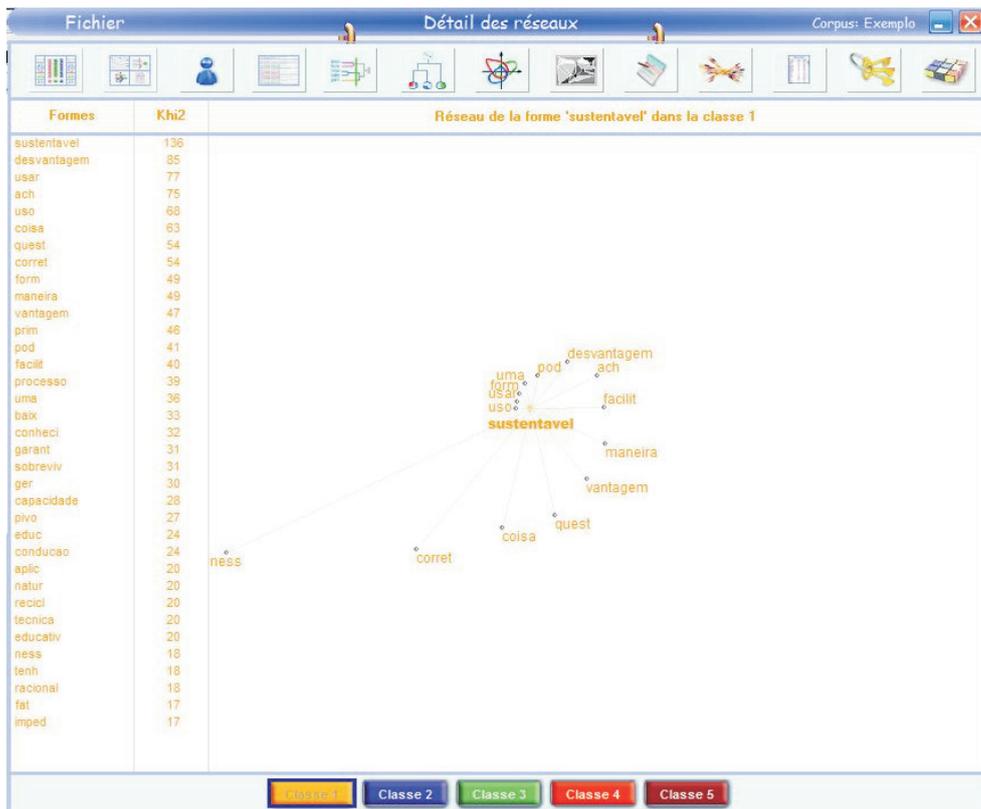
**Comentário:** classificação descendente dos léxicos por classes. Nota-se que, na classe 3, “irrig” aparece próximo de “sistema”, mas ambos distantes de “economia”.

10.5. *Analyse factorielle* – Análise fatorial.



**Comentário:** resultado da análise fatorial, mostrando os discursos dos sujeitos agrupados em classes, podendo-se identificar o sujeito com discurso representativo de cada uma.

10.6. *Détail des Réseaux* – Redes de varejo.



**Comentário:** mostra a aproximação ou distanciamento dos léxicos de uma dada classe com relação ao léxico mais representativo.

### 10.7. Corpus.

The screenshot shows the Alceste software interface with the window title 'Fichier' and 'Corpus'. The main area displays a list of text units (n° uce) with highlighted words. The text is as follows:

1 O nosso #trabalho em relacao ao uso #da agua e com #os #agricultores #familiares. O #foco do nosso\$  
2 trabalho. aqui nos temos a #maior area irrigada do #distrito\_federal, ta concentrada aqui. E essa\$  
3 maior, essa expressao e #devido ao #grande #numero de pivos centrais. esses agricultores\$  
4 #desenvolvem suas atividades em nivel de #organizacao #familiar e alguns e outros individuais. tem\$  
5 #grupos #familiares que #desenvolvem suas atividades, principalmente o seis, ali na #regiao #da area c\$  
6 agora, #os pequenos #produtores, que e o nosso #foco, eles nao tem nenhuma forma de #organizacao\$  
7 nos estamos trabalhando com isso, temos dois #canais na nossa regio de #distribuicao de agua\$  
8 privados e barriamentos. um aqui no #corrego capao\_do\_lu e dois #canais la no #corrego #do #bunit\$  
9 #vermelho, aonde o lineu desenvolve o trabalho dele com #pequenos #reservatorios. E um aqui no\$  
10 #jardim que e la #do #corrego capao\_do\_lu. sao #canais #revestidos. maria #do camo e nossa\$  
11 economista domestica, ja andaram la. E que que acontece, sao pequenos produtores de dificil\$  
12 mento, e #capacidade #natural foram excluidos do #processo de ensino e aprendizagem. voce tem\$  
13 que #considerar isso, mas a tendencia e que eles se organizem, sobretudo pela #escassez da #agua. A\$  
14 #gente ta #passando ai por uma #condicao climatica muito instavel. E, provavelmente essa\$  
15 organizacao #vai vim. esses dois canais, vale #lembrar, que sao #obras estaduais, o governo do\$  
16 #distrito\_federal e o #dono dessas #obras. sao canais publicos, #digamos assim. sao canais publicos\$  
17 de uso #comunitario. existem grupos que domina o uso desses sistemas de irrigacao, desses #canais\$  
18 tem um #grupo de #usuarios. ai, a agua chegou. so se #fosse congelada, ai, pica elas assim. tem\$  
19 #grupos de #usuarios que #dominam o uso #da agua e eles realmente fiscalizam o uso indevido, mas\$  
20 o uso indevido por #pessoas, nao e o uso indevido da #agua, #digamos assim. se uma #pessoa que nao\$  
21 #participa do #grupo usar, eles impedem que eles #continuem usando, tem essa vigilancia. essa\$  
22 vigilancia e so, para-que pessoas fora do #grupo nao usem. quanto ao uso #da agua, nao existe\$  
23 nenhuma preocupacao por #parte deles, e de fato, por exemplo, nos temos dois #periodos distintos\$  
24 de crise de agua. tem um #grupo que #domina, mas nao tem #organizacao. #domina a agua enquanto\$  
25 enquanto #pessoas que podem usar a #agua, so isso, da onde vem. dos pequenos agricultores, e do\$  
26 geral. #dos pequeno, pode colocar ai, todos da #bacia #do #rio\_preto, nossa #area de atuacao na\$  
27 #bacia #do #preto. agora, #dos #pequenos agricultores, sao de #corregos de baixa #vazao. sao #corregos\$  
28 #do #bunit #vermelho. grandes produtores sao o #rio\_jardim, #corrego sao\_bernardo\$  
29 sao\_bernardo, #corrego estreito, e o #rio\_preto. #distribuicao. quando voce fala #comunidade, a\$  
30 gente #entende que e de #uso comunitario. agua de #uso comunitario. que usam comunitariamente\$  
31 #os grandes #produtores estariam fora. #repete. na questao #da #organizacao de #usuario de agua, no\$  
32 #manejo da #irrigacao, no acumulo dessa agua a nivel de #propriedade em reservatorios #eficientes\$  
33 nao tem como usar. E nos estamos #trabalhando para-que isso se foque do nosso #trabalho. manejo\$  
34 #da agua e armazenamento #da agua na propriedade. E acumulo ou armazenamento. sim, #os\$

At the bottom, there are buttons for 'Classe 1', 'Classe 2', 'Classe 3', 'Classe 4', 'Classe 5', and 'Global'.

**Comentário:** interpreta o corpus completo classificando o discurso em classes.

### 10.8. Segments répétés du corpus – Segmentos do corpus repetidos.

Segments répétés	Effectif	Unités de contexte élémentaires
de agua	78	**** "suj_01 "sex_mas "tem_10a "esc_med "exp_ges
a gente	75	O nosso trabalho em relacao ao uso da agua e com os agricultores familiares. O foco do nosso
por exemplo	74	trabalho, aqui nos temos a maior area irrigada do distrito, federal, ta concentrada aqui. E essa
nao tem	50	maior, essa expressao e devido ao grande numero de pivos centrais, esses agricultores
ach que	48	desenvolvem suas atividades em nivel de organizacao familiar e alguns e outros individuais. tem
da agua	48	grupos familiares que desenvolvem suas atividades, principalmente o seis, ali na regio da area c
isso ai	47	agora, os pequenos produtores, que e o nosso foco, eles nao tem nenhuma forma de organizacao
e a	46	nos estamos trabalhando com isso, temos dois canais na nossa regio de distribuicao de agua
a agua	40	privados e barramentos. um aqui no correjo capao_do_ju e dois canais la no correjo do buriti
e o	39	vermelho, aonde o lineu desenvolve o trabalho dele com pequenos reservatorios. E um aqui no
no caso	39	jardim que e la do correjo capao_do_ju, sao canais revestidos, maria do camo e nossa
eu ach que	37	economista domestica, ja andaram la. E que que acontece, sao pequenos produtores de dificil
as vez	35	mento, e capacidade natural, foram excluidos do processo de ensino e aprendizagem, voce tem
e uma	32	que considero isso, mas a tendencia e que eles se organizem, sobretudo pela escassez da agua. A
o uso	32	gente ta passando ai por uma condicao climatica muito instavel. E, provavelmente essa
comec a	31	organizacao vai vim, esses dois canais, vale lembrar, que sao obras estaduais, o governo do
que ele	31	distrito, federal e o dono dessas obras, sao canais publicos, digamos assim, sao canais publicos
que e	30	de uso comunitario, existem grupos que domina o uso desses sistemas de irrigacao, desses canais
tem uma	29	tem um grupo de usuarios, ai, a agua chegou, so se fosse congelada, ai, pica elas assim, tem
e um	27	grupos de usuarios que dominam o uso da agua e eles realmente fiscalizam o uso indevido, mas
de irrig	26	o uso indevido por pessoas, nao e o uso indevido da agua, digamos assim, se uma pessoa que nao
nos temos	26	participa do grupo usar, eles impedem que eles continuem usando, tem essa vigilancia, essa
por cento	26	vigilancia e so, para-que pessoas fora do grupo nao usem, quanto ao uso da agua, nao existe
nao e	25	nenhuma preocupacao por parte deles, e de fato, por exemplo, nos temos dois periodos distintos
eu ach	25	de crise de agua, tem um grupo que domina, mas nao tem organizacao, domina a agua enquanto
que eles	25	enquanto pessoas que podem usar a agua, so isso, da onde vem, dos pequenos agricultores, e do
que a	24	geral, dos pequeno, pode colocar ai, todos da bacia do rio_preto, nossa area de atuacao na
tem um	24	bacia do preto, agora, dos pequenos agricultores, sao de correjos de baixa vazao, sao correjos
na regi	24	do buriti vermelho, grandes produtores sao o rio_jardim, correjo sao_bernardo
faz a	23	sao_bernardo, correjo estreito, e o rio_preto, distribuicao, quando voce fala comunidade, a
do can	23	gente entende que e de uso comunitario, agua de uso comunitario, que usam comunitariamente
o pesso	23	os grandes produtores estariam fora, repete, na questao da organizacao de usuario de agua, no
que tem	23	manejo da irrigacao, no acumulo dessa agua a nivel de propriedade em reservatorios eficientes
	23	nao tem como usar. E nos estamos trabalhando para-que isso se foque do nosso trabalho, manejo
	23	da agua e armazenamento da agua na propriedade. E acumulo ou armazenamento, sim, os
	23	pequenos produtores sao canalizadas, nao, gravidade, sistemas de distribuicao por gravidade em
	23	tubulacao de concreto, ta dando problema pontuais la no buriti vermelho, por causa da ma
	23	qualidade da, do material, isso, em um dos sistemas, irrigacao, hortalias em-geral, hortalias e
	23	frutiferas. E pra completar, seriam as olerícolas, ou nem existe essa palavra no dicionario
	23	olerícolas, olerícolas que sao as plantas da olericultura, que e a area da horticultura que-se dedica
	23	ao estudo da olericultura. E complicado, mas a gente tem, assim da pra entender, nao, essa ideia
	23	num temos ideia do uso da agua porque eles nao tem o manejo, entao a gente nao tem, uma
	23	estimativa ai e muito dificil fazer, por-que eles se atentam pela reducao da irrigacao, por que

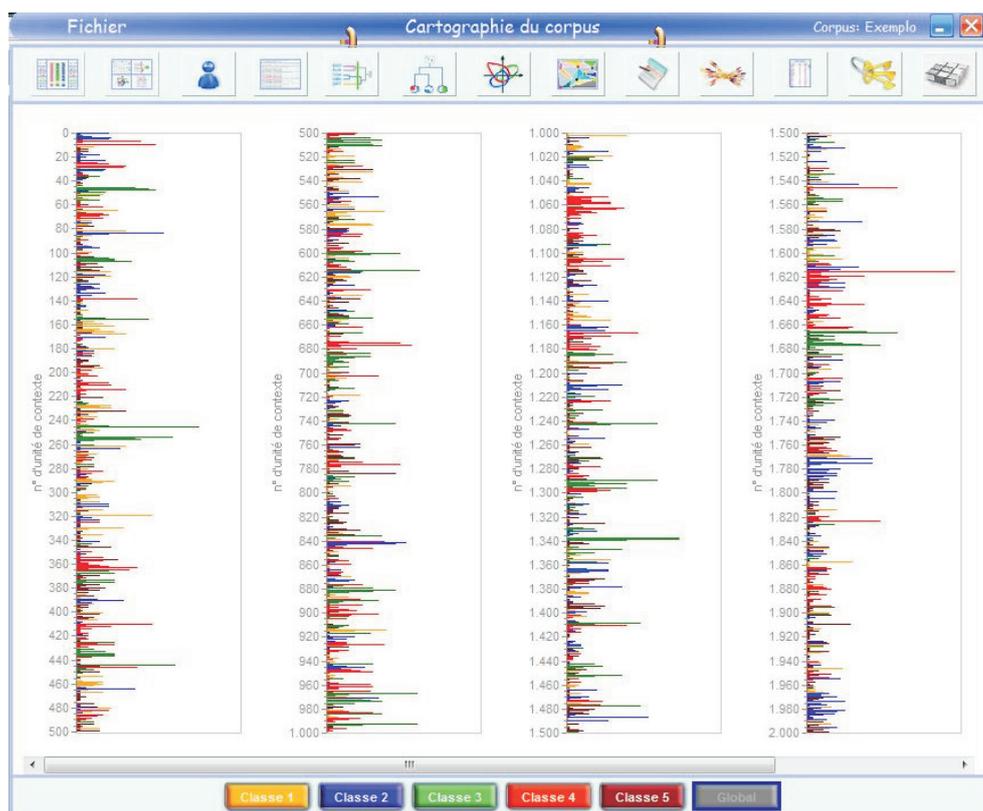
Comentário: são segmentos do corpus repetidos, que tem pouca efetividade na análise. Nota-se que “que tem” é um termo comum, que não foi classificado.

10.9. Dictionnaire du corpus – Dicionário do corpus.

Formes	Effectif dans le corpus	Catégories grammaticales
agua	568	Formes non reconnues
da	547	Formes non reconnues
pra	512	Formes non reconnues
uma	452	Formes non reconnues
do	438	Formes non reconnues
gente	263	Formes non reconnues
irrig	261	Formes non reconnues
vai	249	Formes non reconnues
ach	235	Formes non reconnues
faz	221	Formes non reconnues
as	214	Formes non reconnues
os	193	Formes non reconnues
uso	190	Formes non reconnues
trabalh	173	Formes non reconnues
can	159	Formes non reconnues
pod	159	Formes non reconnues
quest	153	Formes non reconnues
pesso	149	Formes non reconnues
produ	149	Formes non reconnues
num	145	Formes non reconnues
coisa	138	Formes non reconnues
are	120	Formes non reconnues
sistema	117	Formes non reconnues
form	114	Formes non reconnues
fal	110	Formes non reconnues
problema	110	Formes non reconnues
dos	104	Formes non reconnues
nascente	102	Formes non reconnues
falt	91	Formes non reconnues
exemplo	88	Formes non reconnues
dess	87	Formes non reconnues
maior	85	Formes non reconnues
das	83	Formes non reconnues
pro	77	Formes non reconnues

Comentário: traz todas as formas encontradas nas classes. Nota-se que a forma “trabalh” aparece e pode significar trabalho, trabalhador. “Água” foi a forma que mais foi efetiva no corpus.

10.10. Cartographie du corpus – Mapeamento do corpus.



**Comentário:** permite avaliar graficamente variáveis como, por exemplo, número de unidades de contexto elementares de cada classe e no resultado global.

## Conceitos Básicos

**Corpus:** é o banco de dados de texto que será analisado.

Unidade de Contexto Inicial – UCI: é a unidade a partir da qual o programa efetuará a fragmentação inicial; podendo ser capítulos de um livro, respostas de entrevistas, artigos de revistas, entre outros. No caso de entrevistas, as UCIs correspondem a cada entrevista realizada levando em consideração as respostas, isto é, o discurso do participante e, desprezando as perguntas utilizadas.

Por exemplo, o fragmento abaixo refere-se à resposta de um técnico da Emater sobre o uso sustentável da água e, caracteriza-se como uma UCI:

*O nosso trabalho em relação ao uso da água é com os agricultores familiares. O foco do nosso trabalho. Aqui nós temos a maior área irrigada do Distrito Federal, tá concentrada aqui. E essa maior, essa expressão é devido ao grande numero de pivôs centrais. Esses agricultores desenvolvem suas atividades em nível de organização familiar e alguns e outros individuais. Tem grupos familiares que desenvolvem suas atividades, principalmente o seis, ali na região da área c. Agora, os pequenos produtores, que é o nosso foco, eles não tem nenhuma forma de organização. Nós estamos trabalhando com isso, temos dois canais na nossa região de distribuição de água.*

Unidade de Contexto Elementar – UCE: é um segmento de um texto, menor fragmento com sentido; enunciados. É a partir da pertença das palavras de um texto a uma UCE, que o programa vai estabelecer as matrizes para a classificação (REINERT, 1986).

A seguir são apresentados exemplos de UCEs (segmentos de um texto) resultantes da fragmentação de um *corpus* pelo Alceste. Cada UCE vem dividida pelo sinal #.

*#utilizam a #agua #pra #abastecimento #dos #animais e #maioria. #poco #profundo  
ninguem #utiliza aqui #pra #irrigacao, #pra #grande #irrigacao. #pra #pequena  
#utilizada #pra #irrigacao ela e #captada #diretamente, #normalmente #diretamente  
#dos #corregos e ai uma #bomba #pra #succao#*

Unidade de Contexto – UC: segmento maior de texto que contextualiza a UCE, ajuda a compreender o sentido da UCE. Refere-se ao agrupamento de UCEs sucessivas dentro de uma mesma UCI.

No exemplo anterior, pode-se observar UCEs divididas pelo programa, contudo para compreender o conteúdo semântico dessas UCEs deve-se ter um conhecimento prévio e profundo do *corpus* a fim de apreender o significado da UCE naquele contexto mais amplo.

Classe: pode ser definida como um agrupamento constituído por várias UCE de vocabulário homogêneo.

Lematização: operação de substituir palavras por uma forma reduzida, que permite ao programa considerar palavras equivalentes com radical comum.

Exemplo de lematização: IRRIG + que equivale às palavras irriga, irrigação, irrigada, irrigam, irrigante.

Análise Fatorial por Correspondência – AFC: cruzamento entre o vocabulário (considerando a frequência de incidência de palavras) e as classes, gerando uma representação gráfica em plano cartesiano, na qual são vistas as oposições entre classes ou formas.

## Preparação do *Corpus*

O *corpus* deverá ser preparado antes de utilizar o software Alceste.

1. Colocar todos os materiais textuais em um único arquivo do Word. Se forem respostas a questões, colocar as respostas de um mesmo indivíduos juntas, sem as questões, num único arquivo de forma separada por linhas de comando.

Exemplo de recorte de *corpus* de uma entrevista sem as perguntas e aplicada a técnicos da Emater/DF sobre o uso sustentável da água na agricultura:

*O nosso trabalho em relação ao uso da água é com os agricultores familiares. O foco do nosso trabalho. Aqui nós temos a maior área irrigada do Distrito\_Federal, tá concentrada aqui. E essa maior, essa expressão é devido ao grande numero de pivôs centrais. Esses agricultores desenvolvem suas atividades em nível de organização familiar e alguns e outros individuais. Tem grupos familiares que desenvolvem suas atividades, principalmente o seis, ali na região da área c...*

2. O nome deste arquivo deverá ser curto, no máximo 7 caracteres, não podendo haver espaço em branco entre letras, nem *underline*.
3. Separar as UCIs com as linhas de comando (linhas estrelas ou linhas asteriscos). Cada entrevista deverá iniciar por uma linha desse tipo. Não se deve deixar espaço entre a linha de comando e o texto. Mas, deve-se deixar um espaço entre o texto e a próxima linha de comando.

As variáveis que compõem a linha de comando devem ser abreviadas e declaradas separadamente, como por exemplo:

suj: sujeito  
sex: sexo  
mas: masculino  
tem: tempo de experiência  
a: anos  
esc: escolaridade  
med: ensino médio  
exp: área de experiência  
ges: gestão

A linha de comando deverá ser escrita da seguinte maneira: quatro asteriscos, um espaço simples, um asterisco, a variável (suj\_01), um espaço simples, um asterisco, a próxima variável e assim por diante.

Exemplo de linha de comando:

```
**** *suj_01 *sex_mas *tem_10a *esc_med *exp_ges
```

4. Não deixar parágrafos no texto. Este deverá estar “corrido”.
5. Se o texto for oriundo de entrevistas, as perguntas devem ser totalmente suprimidas, ficando apenas as respostas dos respondentes formando a UCI.
6. O texto deverá ser corrigido para que os erros de digitação ou outros não sejam tratados como palavras diferentes. Pode-se usar o recurso do próprio Word.
7. Deve-se observar se há uniformidade com relação às siglas durante o texto.
8. Quando houver palavras compostas, estas deverão ser reunidas por *underline*.

Exemplo: Distrito Federal = Distrito\_Federal

9. Onde houver hífen na palavra, deve-se trocar por *underline*.
10. Não se deve usar aspas, apóstrofes, cifrão, porcentagem, asterisco, interrogação, exclamação.
11. A pontuação deverá ser observada. Os sinais usados são basicamente: ponto (.), vírgula (,) e ponto e vírgula (;).
12. Não se deve usar negrito, itálico, sublinhado ou qualquer outro recurso.
13. Não se deve justificar o texto.
14. Caso haja necessidade de analisar termos compostos enquanto única noção, eles devem estar ligados pelo *underline*.

## Possibilidades de Análise e seus Produtos

O programa faz a composição lexical observando-se as relações entre as palavras. Depois, forma a estrutura temática (classes ou categorias) com base na relação entre as classes: proximidade ou distanciamento entre elas. Pela análise fatorial por correspondência (AFC), ocorre a representação gráfica das posições das classes. Essas estruturas estão estreitamente ligadas à distribuição das palavras em um texto, que raramente ocorre de forma aleatória.

O produto da análise do Alceste refere-se a uma representação icônica do objeto de estudo, o qual orientará a compreensão e a construção de sentido que está subjacente ao processo interpretativo

(JODELET, 1989, citado por ALBA, 2004). Apesar de o material discursivo se tornar diretamente observável, a sua organização latente é sempre objeto de uma reorganização do pesquisador (JODELET, 1989 citado por ALBA, 2004)

## Etapas de Análise Realizadas pelo Software

As etapas de análise realizadas pelo Alceste abrangem:

A: leitura do texto e cálculo dos dicionários.

B: cálculo das matrizes de dados e classificação das UCEs.

C: descrição das classes de UCEs.

D: cálculos complementares.

Na etapa A, ocorre a reformatação e divisão do texto em segmentos (UCEs); a pesquisa do vocabulário e redução das palavras com base em suas raízes (forma reduzida). Ex: irrigar, irrigação, irrigante, irrigam = IRRIG + . Também cria-se o dicionário de formas reduzidas.

Na etapa B, ocorre a seleção das UCEs a serem consideradas e cálculo da matriz: formas reduzidas x UCEs; cálculo das matrizes de dados para classificação hierárquica descendente (CHD) e classificação hierárquica descendente propriamente dita.

Na etapa C, ocorre a definição das classes escolhidas; a descrição das classes e a análise fatorial por correspondência.

Na etapa D, ocorrem os cálculos complementares que abrangem a seleção das UCEs mais características de cada classe; a pesquisa dos segmentos repetidos por classe; a classificação hierárquica ascendente (focada no léxico) e a seleção de palavras mais características de cada classe.

As operações mais importantes para a interpretação do *corpus* são: a classificação hierárquica descendente; a descrição das classes, com radicais mais importantes; e a seleção de UCEs mais características de cada classe.

## Considerações Finais

O Alceste pode agilizar a análise de dados textuais e tem potencial para ser usado no mapeamento geral do *corpus*, resultando em uma classificação objetiva, a partir de critérios estatísticos/matemáticos, baseada na co-ocorrência das palavras; apresentando as relações entre as classes encontradas e gráficos que representam essas relações.

Pode ser usado em conjunto com a análise de conteúdo clássica confirmando ou infirmando tendências.

É um procedimento rápido, podendo ser usado para um grande volume de dados textuais, geralmente oriundo de pesquisas na área social, econômica e ambiental. Com o *corpus* bem formatado, o programa efetua a análise em minutos gerando resultados de melhor qualidade, mais confiáveis.

O Alceste poderá ser usado nos estudos de avaliação da adoção e impacto das tecnologias geradas pela Embrapa.

## Referências

- ALCESTE 2009 : version window : un logiciel d'analyse de données textuelles : manuel d'utilisation. Image, 2009.
- ALBA, M. El método ALCESTE y su aplicación al estudio de las representaciones sociales del espacio urbano: el caso de la ciudad de México. **Paper on Social Representations**, v. 13, n. 1, p. 01-20, 2004.
- BARDIN, L. **Análise de conteúdo**. 3. ed. Lisboa: Edições 70, 2004.
- Camargo, B. V. ALCESTE: um programa informático de análise quantitativa de dados textuais. In: MOREIRA, A. S. P.; JESUÍNO, J. C. (Org.). **Perspectivas teórico-metodológicas em representações sociais**. João Pessoa: Editora Universitária, 2005. p. 511-539.
- CANO, I. **Introdução à avaliação de programas sociais**. 2. ed. Rio de Janeiro: Editora Fundação Getúlio Vargas, 2004.
- COZBY, P. C. **Métodos de pesquisa em ciências do comportamento**. São Paulo: Atlas, 2003.
- DENZIN, N.; LINCOLN, Y. (Org.). **O planejamento da pesquisa qualitativa: teorias e abordagens**. Porto Alegre, Bookman, 2006.
- JODELET, D. (Org.). **Les représentations sociales**. Paris: PUF, 1989. 449 p.
- Software ALCESTE: analyse lexicale par context d'un ensemble de segments de texte, Image, 2009.
- MARTINS, C. R. **Sistema Único de Saúde: uma análise das crenças dos seus representantes em municípios rurais na Paraíba**. 2008. Dissertação (Mestrado). Universidade Federal da Paraíba. Programa de Pós-Graduação em Psicologia Social. João Pessoa, PB.
- MINAYO, M. C. S. **O desafio do conhecimento: pesquisa qualitativa em saúde**. 4. ed. São Paulo: Hucitec; 1996. 269 p.
- MINAYO, M.C.S. (Org.). **Pesquisa social: teoria, método e criatividade**. 13. ed. Petrópolis: Vozes, 1999.
- NASCIMENTO, A. R. A.; MENANDRO, P. R. M. Análise lexical e análise de conteúdo: uma proposta de utilização conjugada. **Estudos e Pesquisa em Psicologia**, v. 6, n. 2, p. 72-88, 2006.
- REINERT, M. Un logiciel d'analyse lexicale: ALCESTE. **Les cahiers de l'Analyse des Données**, v. 4, p. 471-484, 1986.
- RICHARDSON, R. J. **Pesquisa Social**. 3. ed. São Paulo: Atlas, 1999. 328 p.
- ROCHA, F. E. de C.; ALBUQUERQUE, F. J. B. de; MARCELINO, M. Q. dos S.; DIAS, M. R.; PINHEIRO, J. de Q. **Aplicação da análise de conteúdo na perspectiva de Bardin em uma aproximação avaliativa do Pronaf-PB**. Planaltina: Embrapa Cerrados, 2008. (Boletim de pesquisa e desenvolvimento, 201).
- SOUZA, M. L. de; GOMES, W. B. Evidência e interpretação em pesquisa: as relações entre qualidades e quantidades. **Psicologia em estudo**, v. 8, n. 2, p. 83-92, 2003.
- TRIVINOS, A. N. S. **Introdução à pesquisa de ciências sociais**. São Paulo: Atlas, 2006.
- ZANELLI, J. C. Pesquisa qualitativa em estudos da gestão de pessoas. **Estudos de psicologia**, v. 7 (spe), p. 79-88, 2002.

## **Software Guide Alceste**

---

### **Abstract**

*This guide presents Alceste software (Analyse lexicale par context d'un ensemble de segments de texte) as a computerized tool for analyzing texts from written and oral communication. The program performs analysis descriptive statistics and mathematics to measure the text and bring out structures that have more significant meanings. Traditionally, the software is applied in the field of linguistics and was introduced in the humanities and social sciences that use texts as a means of preserving the quality of the phenomena studied. However, the software can be used in the evaluation of social programs. At Embrapa, could be used in assessing the adoption of technologies generated by the company, in technical operations and the lifting of indicators to assess impact / results focused on people's opinions.*

*Index terms: text analyse, lexical analyse, Alceste software.*