

APRESENTAÇÃO DO DICIONÁRIO BILINGUE CABOVERDIANO-PORTUGUÊS

NICOLAS QUINT
(CNRS/LLACAN)

MAFALDA MENDES
(Verbalis)

FÁTIMA RAGAGELES
(Verbalis)

Convenções Gráficas

mágua: palavra crioula caboverdiana.

mágoa: palavra portuguesa.

"jaloux": palavra francesa.

|ciumento|: sentido da palavra.

/magwe/: transcrição fonológica,

{tch}: grafema.

Introdução

O *Dicionário Caboverdiano-Português*, que aqui estamos a apresentar, é o fruto da minha colaboração, como linguista especializado na área da linguística caboverdiana, com a Verbalis, empresa que trabalha no desenvolvimento de instrumentos linguísticos com base informática para a língua portuguesa. Também contamos com a participação da Priberam Informática, que contribuiu decisivamente para a produção dos suportes informáticos do dicionário, e com a consultoria linguística da Dra. Dulce Pereira, especialista portuguesa do crioulo caboverdiano.

A nossa apresentação constará de três partes:

- primeiro, eu falarei da minha tarefa de recolha da informação lexicográfica que desembocou na elaboração dum *corpus* normalizado de palavras caboverdianas;

- depois, a Dra. Mafalda Mendes, da Verbalis, nos contará como efectuou o processo de adaptação dos dados recolhidos às tecnologias informáticas mais modernas de tratamento das línguas, e as consequências e vantagens de tal tratamento.

- por fim, Fátima Ragageles, da Verbalis, explicar-nos-á quais foram as dificuldades técnicas de tradução com que teve de se enfrentar para chegar à elaboração de uma base de dados Caboverdiano-Português, já que eu inicialmente descrevera a língua caboverdiana a partir da minha língua materna, o Francês.

Informação lexicográfica e metodologia de trabalho

Conteúdo da obra

O nosso dicionário de Caboverdiano-Português é um dicionário bilingue que proporciona a tradução sistemática de mais de 4.000 palavras caboverdianas para o Português. Publicado em 1998 pela Verbalis, foi apresentado por primeira vez quando da Expo 98 no pavilhão da República de Cabo Verde. Trata-se do único dicionário Caboverdiano-Português actualmente disponível no mercado. Existe em formato papel e CD-ROM e, em 1999, foi inaugurada a versão on-line, acessível no site <http://www.priberam.pt/dcvpo>.

Para além da informação lexical nele contida, esse dicionário também propõe etimologias portuguesas, africanas ou outras, para mais de 90% das palavras contempladas.

A variante linguística que serve de referência é o crioulo da ilha de Santiago, falado pela maioria da população caboverdiana.

Recolha de informação

Até hoje, a língua oficial da República de Cabo Verde é o Português. A língua de Camões e de Pessoa é utilizada em todos os documentos oficiais, na escola e na imprensa. Não obstante, enquanto falante, o conjunto da população caboverdiana expressa-se em crioulo caboverdiano, ou língua caboverdiana.

Quando cheguei a Cabo Verde, em 1995, depressa decidi aprender a língua crioula, para poder conversar mais livremente com os caboverdianos. Fui a várias livrarias da Praia, em busca de algum dicionário ou léxico de crioulo que me pudesse ajudar, mas nessa altura não havia nenhum à venda¹. Comecei logo a constituir o meu próprio dicionário, a partir das conversas espontâneas dos caboverdianos que me rodeavam. Cada vez que ouvia uma nova palavra crioula, anotava-a. Também anotava a frase ou o contexto em que fora proferida e muitas vezes pedia aos meus amigos caboverdianos esclarecimentos sobre o significado da palavra, o nível de língua implicado, a frequência de uso, etc. Andava sempre

com uma folha de papel e um lápis no bolso para escrever todas essas informações e, todos os dias, registava o conteúdo do papel no computador. Dessa maneira, pouco a pouco, fui constituindo o meu *corpus* de Caboverdiano.

Mas não bastava recolher a informação. Também foi preciso escolher. Na Praia, convivem pessoas oriundas de várias ilhas de Cabo Verde, e cada qual fala com o seu sotaque, com as suas palavras.

Exemplos:

- a palavra *vez* corresponde a *bés* /bes/ em crioulo de Santiago, mas a *bes* /bes/ em crioulo do Fogo

- a palavra *carro* corresponde a *káru* / 'káru em crioulo de Santiago, mas a *kór* /kɔɾ/ em crioulo do Mindelo. E os próprios santiaguenses, os chamados badios, também não falam todos da mesma maneira. Os moradores da Praia incorporam várias palavras provindas do Português moderno no seu idioma, enquanto os habitantes do interior da ilha falam uma variante mais original, o chamado crioulo fundo.

Exemplo:

- a palavra *balancear* corresponde a *balansia* /be'lãsɿə/ na Praia, mas a *djolongi* /dʒo'lõgi/ em crioulo fundo do interior.

Para preservar a coerência do meu *corpus*, escolhi então privilegiar o crioulo fundo de Santiago. Escolhi o crioulo de Santiago por duas razões:

1. quando constituí o *corpus*, eu vivia na ilha de Santiago e tinha um contacto quotidiano com essa variante da língua caboverdiana;

2. mais da metade da população caboverdiana tem o santiaguense como língua materna. Um léxico ou dicionário de santiaguense teria mais utilidade para os estrangeiros (como eu) e até para os próprios caboverdianos.

Escolhi o crioulo fundo por outras duas razões:

1. dum ponto de vista patrimonial ou cultural, o crioulo fundo é a variedade mais genuína da língua caboverdiana, relativamente ao Português. Registrar o seu vocabulário e as suas expressões mais típicas seria um contributo para a preservação e valorização da identidade caboverdiana;

2. dum ponto de vista didáctico e prático, depressa me dei conta de que era mais fácil aprender o santiaguense rural precisamente porque este se diferencia mais do Português. Efectivamente, embora a proximidade lexical do Português e das variedades crioulas mais aportuguesadas ajude às vezes à compreensão (para um lusofalante, é mais fácil compreender o sentido de *balansia* do que o de *djolongi*), também favorece as confusões entre ambas as línguas. E, de facto, esse perigo de confusão é uma fonte de problemas para todos em Cabo Verde:

- para os próprios caboverdianos, que muitas vezes julgam que todas as palavras e expressões da sua língua podem ser transpostas exactamente para o Português, o que não é o caso. Assim, um caboverdiano, amigo meu, que nessa altura estudava no liceu da Praia, contou-me ter ouvido um colega que, expressando-se em Português, tinha dito ao professor que a resposta estava *mareada*. O verbo crioulo *maria* /'mɛrjɐ/ parece-se muito com o português *marear*, mas em português, as respostas não marciam, *estão erradas*.

- para os estrangeiros vivendo em Cabo Verde, os quais, na maioria dos casos lusofalantes ou com algum conhecimento da língua portuguesa, pensam que para falar bem crioulo, só é preciso modificar um pouco as desinências das palavras portuguesas. Mas infelizmente (ou felizmente, porque cada língua tem o seu encanto) não é tão fácil falar correntemente o crioulo, mesmo que se trate de uma variedade mais lusitanizada.

Por conseguinte, o crioulo fundo de Santiago pareceu-me uma variedade a privilegiar, porque concentra a identidade linguística caboverdiana e ao mesmo tempo favorece a tomada de consciência das diferenças profundas que separam as línguas crioula e portuguesa.

Formalização dos dados

Tendo escolhido uma variedade determinada, tinha ainda de escolher convenções coerentes de transcrição dos meus levantamentos. Decidi transcrever cada palavra recolhida sob duas formas:

1. a transcrição fonética. Cada palavra do *corpus* foi transcrita segundo as convenções do alfabeto fonético internacional, o que garante o rigor do trabalho e favorece a sua exploração pelos linguistas e científicos;

2. uma transcrição ortográfica. De facto, a maioria dos utentes potenciais dum futuro dicionário de crioulo, quer sejam caboverdianos ou estrangeiros, para desfrutar correctamente da informação lexicográfica, precisa de uma grafia menos alheia aos seus hábitos de leitura do que a transcrição fonética. Em 1995 e 1996, quando estive a fazer os meus levantamentos, ainda não existia nenhuma ortografia de referência do Caboverdiano, mas já havia tendências convergentes, como por exemplo o uso sistemático da letra {k} pela maioria dos autores crioulófonos para transcrever o fonema /k/. Elaborei um sistema ortográfico da língua caboverdiana preocupando-me sempre em:

- produzir uma grafia fonemática que fizesse corresponder cada fonema a determinado signo (ou grupo de signos) gráfico(s). Assim, escolhi distinguir graficamente {á} aberta /a/ e {a} fechada /ɐ/, já que estes sons permitem distinguir pares minimais em crioulo, como *mágua* /'magwɐ/ *mágoa* ≠ *magua* /'mɛgwe/ *magoar*;

- privilegiar as grafias consensuais quando existiam (como por exemplo o uso do {k}, baseando-me nos hábitos dos linguistas e pesquisadores que me

antecederam, em particular nos trabalhos do linguista caboverdiano Manuel Veiga;

- favorecer as soluções que respeitassem os hábitos gráficos do povo caboverdiano, habituado às convenções ortográficas portuguesas. Foi por isso que, por exemplo, transcrevi o adjectivo crioulo **bom** /bõ/ *bom* como em português (já que todos os caboverdianos aprenderam na escola a grafar em Português as vogais nasais finais com um *-{m}*), e não **bon**, forma preferida por outros linguistas;

- ter em conta as criações espontâneas dos caboverdianos, dando especial atenção aos *graffitti*, testemunhas do génio popular crioulo. Foi assim que preferi transcrever a africada /t / por {tch} (outros preferem escrever {tx}) seguindo o que me sugeriram numerosos escritores caboverdianos anónimos nas paredes da sua terra.

A Base de dados lexicográfica

O *Dicionário Caboverdiano-Português, Variante de Santiago*, assenta numa base de dados lexicográfica em suporte *Microsoft Access*. Esta base de dados foi desenhada e desenvolvida na Verbalis a partir dos dados lexicais bilingues Caboverdiano-Francês fornecidos por Nicolas Quint em suporte *Microsoft Word*. O projecto da Verbalis para desenvolvimento de um dicionário electrónico Caboverdiano-Português teve como objectivo último dotar a linguística caboverdiana com alguns dos recursos tecnológicos hoje disponíveis para o tratamento das línguas naturais, os quais se podem revelar fundamentais para o apetrechamento da língua caboverdiana, em tempo mais ou menos útil, com os instrumentos linguísticos básicos (léxicos, dicionários, gramáticas, *corpora* de referência da língua caboverdiana) para a sustentação da sua vitalidade como língua veicular de conhecimento e cultura em todas as situações de comunicação de uma sociedade moderna. Numa língua em que quase tudo está por fazer, do ponto de vista da sua descrição e normalização, importa lançar as mãos a essa tarefa recorrendo às metodologias e tecnologias mais produtivas e eficazes.

A base de dados lexicográfica desenvolvida não se pretende como um objecto estático e perfeito, nem se esgota no *Dicionário Caboverdiano Português*. A base de dados constituída pretende-se uma plataforma de trabalho fiável e dinâmica que permita a continuidade e evolução do trabalho lexicográfico sobre a língua caboverdiana, o qual foi apenas aflorado nesta obra bilingue.

Concepção e desenvolvimento da base de dados lexicográfica

A base de dados foi desenhada tendo em conta a micro-estrutura implícita no texto dos verbetes Caboverdiano-Francês fornecidos pelo autor. O trabalho lexicográfico que nos chegou às mãos era aparentemente consistente e

sistemático no que respeita à escrita dos verbetes, permitindo inferir com alguma segurança a micro-estrutura subjacente à escrita dos verbetes e permitindo efectuar uma migração automática de dados de um suporte para o outro. O resultado da migração da informação, porém, possibilitou a detecção de algumas inconsistências na formalização dos dados em formato de texto, aliás típicas do trabalho lexicográfico em papel, e acabou por se revelar uma eficaz ferramenta de revisão do acervo preexistente. Este género de problemas é causador de ruído no processo de descodificação da informação pelo consulente humano e é-o também no processamento automático da informação.

Poderíamos talvez distinguir duas grandes áreas de problemas detectados. Por um lado, flutuações a nível das convenções gráficas para notação dos diferentes tipos de informação lexical veiculada: uso dos tipos negrito, itálico e normal; uso dos diversos tipos de parênteses, flutuação na selecção das abreviaturas, etc. A outro nível, situam-se problemas algo mais profundos e de solução menos evidente, que se prendem com a homogeneidade da aplicação dos critérios usados para discriminação das diferentes camadas de informação lexicográfica em cada verbete ou com a própria conceptualização da informação. Veja-se, a título de exemplo, os seguintes verbetes:

I - exemplos de verbetes do original Caboverdiano-Francês

Ndoxa [ˈ~doʃa] v (E. port. *en-* + *doce* + *-ar*) 1. sucrer, 2. manger des choses sucrées, **ndoxa pa-u ba deta**, prends de la "confiture" (de noix de coco) avant d'aller te coucher.

É [e] v (E. port. (*ele*) *é, de ser*) 1. Forme présente du verbe *sér*, être, v. **éra, sér**, a. sens de *ser* en portugais, **el é bazófu**, il est beau, élégant, **bandidu é bo!**, bandit toi-même!, b. sens de *estar*, **el é la**, il est là-bas, v. **sta**, 2. **sér pa**, falloir, **é pa-u fia-s**, il faut que tu (le) leur dises (...)

No primeiro verbete relativo à palavra **ndoxa** a numeração árabe tem como função diferenciar dois universos semânticos denotados por **ndoxa**. No segundo verbete, já a numeração árabe não parece cumprir essa função e do ponto de vista conceptual a informação discriminada por 1. e 2. não parece ser do mesmo tipo. Por seu turno, surgem as alíneas a. e b., as quais parecem assumir a função que no verbete anterior era desempenhada pela numeração árabe.

Estes casos deram origem a um frutuoso diálogo entre o autor do dicionário e a equipa de trabalho da Verbalis, ao longo do qual foi possível trazer mais alguma luz a zonas menos claras do texto original. Note-se que o trabalho lexicográfico tradicional, baseado em fichas avulsas de informação, em última análise perpetuado no trabalho em ficheiros de texto electrónico, é naturalmente

propício a este tipo de deslizes conceptuais, de difícil detecção em suportes lineares de grandes acervos de informação textual.

Estrutura da base de dados lexicográfica trilingue Caboverdiano-Francês-Português

A base de dados lexicográfica produzida assenta em quatro tabelas de informação.

II - tabelas de informação
tabela de entradas - [Entradas]
tabela de categorias - [Categorias]
tabela de sentidos - [Sentidos]
tabela de exemplos - [Exemplos]

Relações entre as tabelas:
[Entrada] 1 - — [Categorias]
[Categorias] 1 - — [Sentidos]
[Sentidos] 1 - — [Exemplos]

O esquema pressupõe uma hierarquia que se expande de cima para baixo e da esquerda para a direita. A tabela de entradas funciona como a tabela topo de uma hierarquia de camadas de informação lexical. Cada entrada pode ramificar-se em várias categorias morfo-sintácticas; para cada categoria morfo-sintáctica pode dar-se o caso de polissemia, e cada sentido, por sua vez, poderá ser ilustrado por vários exemplos de uso ou lexias complexas.

Os campos que constituem as diferentes tabelas discriminam diferentes níveis de informação lexical. Nalguns casos a discriminação efectuada é ainda incipiente ou relativamente pouco fina, o que se prende com o carácter pioneiro e ainda exploratório deste trabalho. Estes campos são marcados na coluna "Estado" pela abreviatura ED que significa "em desenvolvimento". Por exemplo, os campos "vd." das tabelas [sentidos] e [exemplos] deverão vir a ser desdobrados, em fases ulteriores de desenvolvimento, em campos para anotação de diferentes relações semânticas, tais como sinonímia, antonímia, hiponímia, etc.

III - Tabelas e respectivos campos de registo de informação

[Entrada]

Campo	Descrição	Estado
vedeta	forma gráfica da palavra de entrada - normalmente um lema	
transcrição fonética	transcrição fonética da palavra de entrada	
etimologiaPT	informação de cariz etimológico veiculada em Português	
etiniologiaFR	informação de cariz etimológico veiculada em Francês	
O_m_q	remissão para variantes da palavra de entrada ou sinónimos puros	ED
morfologiaPT	informação morfológica sobre a palavra de entrada veiculada em Português	
morfologiaFR	informação morfológica sobre a palavra de entrada veiculada em Francês	

[Categorias]

Campo	Descrição	Estado
categoriaPT	categoria morfo-sintáctica (segundo a nomenclatura gramatical portuguesa adoptada pela Verbalis)	
categoriaFR	categoria morfo-sintáctica (segundo a nomenclatura gramatical francesa adoptada pelo autor)	
morfologiaPT	informação sobre possibilidades de flexão morfológica veiculada em Português	
morfologiaFR	informação sobre possibilidades de flexão morfológica veiculada em Francês	

[Sentidos]

Campo	Descrição	Estado
sentidoPT	equivalente ou definição do sentido em Português	
sentidoFR	equivalente ou definição do sentido em Francês	
dom.	área temática	ED
constr.	informação de teor colocacional ou lexicas complexas com base na palavra de entrada	ED
vd.	remissões de natureza semântica	ED

[Exemplos]

Campo	Descrição	
exemplo	frase exemplificativa do uso da palavra de entrada ou fraseologia associada à palavra de entrada	
traduçãoPT	equivalente ou definição do sentido em Português	
traduçãoFR	equivalente ou definição do sentido em Francês	
sentido literalPT	sentido literal em Português	
sentido literalFR	sentido literal em Francês	
dom.	área temática	ED
vd.	remissões de natureza semântica	ED

O acervo de informação linguística

Passamos a descrever o perfil do acervo de informação linguística disponível na base de dados Caboverdiano-Francês-Português que poderá dar uma imagem das potencialidades do mesmo para futuros trabalhos:

- registo de uma proposta de forma gráfica para 4222 vocábulos do Caboverdiano,
- registo de uma proposta de transcrição fonética para todos os vocábulos de entrada,
- categorização morfo-sintáctica para cada um dos 4222 vocábulos de entrada;
- registo de 4732 frases e expressões do Caboverdiano que resultaram da transcrição de produções linguísticas atestadas por Nicolas Quint na ilha de Santiago, as quais constituem um pequeno mas valioso *corpus* de referência para a língua caboverdiana tal como ela é falada actualmente nas regiões rurais da ilha de Santiago;
- às 4732 frases corresponde um total de 19077 palavras do Caboverdiano (incluindo palavras repetidas) em que se diferenciam 3727 formas lexicais. Destas 3727 formas lexicais, 1624 não correspondem às formas gráficas dos vocábulos de entrada do dicionário, pelo que deverão corresponder a formas flexionadas dos mesmos ou a palavras não contempladas na nomenclatura;
- cada vocábulo caboverdiano de entrada na base de dados está duplamente associado a termos de sentido equivalente em Português e a termos de sentido equivalente em Francês ou a definições de sentido redigidas em Português e a definições de sentido redigidas em Francês (ou seja, trata-se de uma base de informação trilingue);

- o mesmo se passa para o conjunto de frases ou expressões do Caboverdiano;
- a partir do campo de registo de equivalências para Português foi possível obter 3480 palavras do Português, as quais correspondem a equivalentes para os sentidos de 2473 vocábulos do Caboverdiano; estas palavras encontram-se relacionadas entre si por relações de sinonímia;
- 923 destas palavras portuguesas encontram-se associadas a frases ou expressões do Português com tradução em Caboverdiano (ex" *defender*| vou defender-te (em tribunal)| *m-ta'a divoga pa bo*);
- registo de étimos propostos para 3803 vocábulos do Caboverdiano;
- registo de informação temática para 2929 das acepções de sentido propostas para os vocábulos de entrada do Caboverdiano (no total foram atribuídas 5207 acepções de sentido diferentes para os 4222 vocábulos de entrada);
- registo de informação temática para 637 das frases ou expressões do Caboverdiano.

Multifuncionalidade e reusabilidade dos dados

O potencial de reusabilidade e de multifuncionalidade da base de dados lexicais constituída faz do acervo de informação acima descrito um caso ímpar no panorama dos estudos linguísticos caboverdianos. Na verdade, poderíamos pensar em outras obras de grande potencial informativo sobre a língua caboverdiana, tal como os trabalhos de Napoleão Fernandes, Baltasar Lopes, Tomé Varela, entre outros, que, devido à rigidez e fechamento do seu suporte informacional, o papel se tornam de difícil e moroso manejo.

Assim, a entroncar nesta base de dados, a Verbalis e Nicolas Quint têm já em agenda uma diversificada carteira de projectos na área da linguística caboverdiana começando pelo dicionário bilingue Português-Caboverdiano.

A importância deste potencial torna-se tanto mais evidente quando perante uma língua para a qual tão poucos recursos existem.

Metodologia da tradução

Processo de fixação dos verbetes em Português

Todo o trabalho de desenvolvimento da versão portuguesa do dicionário foi realizado na base de dados criada para o efeito.

Ao longo da produção da primeira versão, esta ia sendo enviada ao autor do dicionário e à consultora na área da linguística caboverdiana, Dra. Dulce Pereira, para uma primeira revisão. O resultado desta revisão era depois inserido na base de dados, seguindo, posteriormente, para a revisão final. Não obstante, sempre que se julgasse necessário, havia revisões intermédias.

Princípios de orientação do trabalho de escrita

Pretendeu-se, ao traduzir a informação da língua de partida para a língua de chegada, dar o equivalente português mais próximo da palavra de entrada caboverdiana ou, na sua ausência, descrições do seu universo de sentidos, usando uma linguagem clara, correcta e corrente, de modo a permitir a boa compreensão da informação dada.

Principais problemas

Ilustra-se em seguida alguns dos problemas de tradução mais significativos.

Falsos amigos

O problema dos falsos amigos pôs-se com frequência, uma vez que se trabalhava em três línguas cujo vocabulário é, basicamente, latino.

Uma palavra ou uma expressão, aparentemente igual, pode ter evoluído de maneiras diferentes em cada uma das línguas, não sendo isso imediatamente perceptível. Tome-se o exemplo da palavra caboverdiana **namora** e da portuguesa *namorar*, a primeira remetendo para uma relação essencialmente física e a segunda para uma relação essencialmente sentimental. Optou-se aqui pela tradução *fazer amor*, equivalente à francesa "faire l'amour", a qual reflecte a relação sensual, sentimental ou não, a que **namora** se refere.

Dá-se também o caso de palavras ou expressões cuja tradução literal, mais do que apontar para um sentido diferente, aponta para um sentido oposto: a expressão caboverdiana **teni algem na korasom** quer dizer em português [sentir rancor por alguém]. Se tomarmos a sua tradução literal teremos a expressão *ter alguém no coração* e vemos que ela quer dizer [sentir amor por alguém]. Optou-se, em casos como estes, por dar também a tradução literal, o que permite a imediata constatação de que a mesma expressão encerra conceitos opostos. Vejamos agora a expressão **bóka báxu**, que quer dizer [de cabeça para baixo, de pernas para o ar, ao contrário]. Ora, em Português, também existe a expressão *de boca para baixo*, a qual aparentemente equivale às dadas como traduções possíveis da expressão caboverdiana. No entanto, em Caboverdiano pode virar-se o que quer que seja **bóka báxu**, assim como em Português se pode virar qualquer coisa *de cabeça para baixo, de pernas para o ar ou ao contrário*, mas apenas algumas coisas podem ser viradas *de boca para baixo*. O universo semântico a que se refere **bóka báxu** é muito mais lato do que aquele a que se refere *de boca para baixo* e esta expressão nunca seria uma boa tradução da anterior.

Polissemia em três línguas

Houve, ainda, a preocupação de delimitar o mais exactamente possível o universo semântico denotado pela palavra ou expressão caboverdiana, o que,

partindo da tradução francesa, se mostrou por vezes difícil. Nalguns casos, a tradução francesa parecia poder dar mais do que um sentido àquilo que se estava a traduzir ou apontar para um universo conceptual algo vago. Noutros casos, pelo contrário, o conceito denotado pela tradução francesa poderia ser mais restritivo do que o denotado pelo vocábulo caboverdiano.

Eis alguns exemplos típicos de situações problemáticas surgidas:

pupa havia sido traduzido por "pousser des cris perçants, aboyer (chien), croasser (corbeau), piailler (oisillon), crier (cochon qu'on égorge...), rejaillir, rugir (eau qui tombe sur des rochers)", não sendo claro se esta palavra denotava o conceito geral de ruído pungente| ou se apenas cada um dos sons especificados na definição francesa;

toku havia sido traduzido por "souche de bananier, tronc coupé resté fiché en terre". Neste caso não era evidente se o toco de que se falava era apenas o da bananeira, ou se podia ser o toco de qualquer árvore, embora especialmente o da bananeira;

sáku havia sido traduzido por "sac en toile, sac à grains". Punha-se a questão de se referir a um qualquer saco de pano e a uma saca para guardar leguminosas secas, ou de se referir apenas a uma saca de pano para guardar leguminosas secas;

orenta havia sido traduzido por "se refaire (une santé financière)". A dúvida que aqui se punha era se a palavra caboverdiana queria dizer especificamente |refazer-se financeiramente| ou se o parêntesis era apenas ilustrativo de um dos casos a que a palavra se poderia referir;

siumentu havia sido traduzido para o Francês por "jaloux" que, querendo dizer |ciumento|, também pode querer dizer |invejoso|. sentido este que, consultando o autor, soubemos que a palavra caboverdiana não contempla. Apesar das parecenças entre as palavras **siumentu** e *ciumento* e apesar de "jaloux" querer quase sempre dizer ciumento, sentiu-se a necessidade de confirmar com o autor que, de facto, apenas este sentido era comportado.

Estratégias de solução dos problemas de tradução

A solução dos problemas surgidos passou, em última instância, pelo recurso sistemático ao autor, à consultora e a informantes caboverdianos e portugueses. Nem todos os problemas ficaram resolvidos de modo completamente satisfatório e nem todas as dúvidas completamente esclarecidas. Ou porque o tempo foi escasso ou porque nem o autor ou a consultora ou mesmo os próprios informantes de língua materna caboverdiana sentiram ter segurança suficiente para responder com precisão às questões levantadas. No entanto, todos os problemas e dúvidas com que nos deparámos se encontram anotados e descritos na base de dados, tendo alguns sido posteriormente resolvidos e sendo

os outros passíveis de futura resolução, o que certamente acontecerá na continuidade de um trabalho que dificilmente se dará por acabado.

Criação da classificação temática

Houve, ao longo do trabalho, a percepção cada vez mais clara de que existiam várias famílias temáticas de maior representatividade dentro do dicionário. Criou-se então uma grelha de classificação que pudesse dar conta desses universos temáticos recorrentes. As palavras e os seus diversos sentidos foram então classificadas de acordo com este novo parâmetro. Deste modo, e utilizando as funcionalidades de pesquisa inerentes à plataforma de trabalho *Microsoft Access*, tornou-se mais fácil gerir e manter a coerência na aplicação dos critérios de tratamento das palavras dentro de determinada área. Procurou-se, por exemplo, garantir que as definições dentro da mesma família temática obedecessem tanto quanto possível à mesma fórmula ou que palavras ou expressões sinónimas, por exemplo, **korótchi** e **porku di tchom**, cuja tradução é *hicho-de-conta*, pudessem ter definições coerentes entre si.

A título de exemplo de uma grande família temática vejamos o caso das palavras e expressões relacionadas com o universo rural de Cabo Verde, cujo acervo é muito importante. Há referências às partilhas e divisões de água e de terras, às alfaías agrícolas, às sementeiras e às colheitas, ao gado e aos animais, de criação, às estações do ano e ao clima. Para que a informação fosse o mais fiável possível, e os termos empregues correntes no vocabulário campesino, recorreu-se a informantes familiarizados com o mundo rural português, tendo a vantagem de a exploração rural em Cabo Verde ser similar àquela que em alguns casos ainda se pratica nos minifúndios portugueses. A possibilidade de ter acesso imediato a toda a área temática, ou a qualquer parte dela, facilitou grandemente o seu tratamento sistemático.

Conclusão

Este dicionário bilingue Caboverdiano-Português representa um esforço importante e concreto de fixação por escrito duma língua que, até hoje, só se manteve por via oral. O nosso trabalho não é um fim, mas um princípio que pode prolongar-se de diversas maneiras:

A. ampliando a base de dados já constituída, porque se as 4.000 palavras recolhidas representam uma boa parte do vocabulário corrente, estas não bastam para as necessidades expressivas duma sociedade moderna;

B. desenvolvendo outras ferramentas para a fixação do crioulo, em particular um dicionário bilingue Português-Caboverdiano, um método de aprendizagem sistemática do crioulo caboverdiano, uma gramática de referência e um livro de textos literários em Caboverdiano;

C. promovendo o uso dum crioulo de referência no ensino em Cabo Verde, bem como nas comunidades caboverdianas emigradas, a fim de favorecer a assunção da língua crioula pelos seus falantes e o estabelecimento de um bilinguismo equilibrado com o Português que, esperamos, substituirá no futuro a presente situação de diglossia vivida por muitos caboverdianos que, muitas vezes, pensam e falam em crioulo e escrevem em Português sem nunca dominar completamente as riquezas das duas línguas de que são detentores.

O facto de os deputados da Assembleia Legislativa da República de Cabo Verde terem votado, em Março de 1998, a oficialização do crioulo como primeira língua nacional, sendo o Português a primeira língua estrangeira, parece-nos criar condições favoráveis para chegar a este futuro bilinguismo. O carácter plurinacional da equipa que produziu este dicionário, em que colaboraram investigadores franceses e portugueses, bem como informantes caboverdianos, é também um ponto positivo. A fixação e a promoção das línguas dos povos que não têm a sorte de poder estudar ou trabalhar no seu idioma materno representam um contributo valioso para o seu desenvolvimento e um objectivo válido, para cujo cumprimento os estudiosos e linguistas devem colaborar, sejam portugueses, caboverdianos ou franceses.

Notas

1 Alguns meses depois um amigo trouxe-me da ilha do Sal o livro de Armando Napoleão Fernandes, *Léxico do dialecto crioulo do arquipélago de Cabo Verde*, Gráfica do Mindelo Lda., Mindelo, 1991. Trata-se de uma obra muito rica e interessante, mas é preciso conhecer muito bem o crioulo para poder disfrutar da informação nela contida, já que apresenta palavras provenientes de diversas ilhas e que usa um sistema ortográfico bastante original.

Bibliografia:

Língua Caboverdiana ou Crioulo de Cabo Verde:

BRITA, A. Paulo de, "Apontamentos para a gramática do crioulo que se fala na ilha de Santiago, Cabo Verde", (1885), in F. Adolfo Coelho "Os dialectos românicos ou neo-latinos na África, Ásia e América" (1881), in *Estudos Linguísticos Crioulos*, Lisboa, ed. Academia Internacional de Cultura Portuguesa, 1967.

FERNANDES, Armando Napoleão, *Léxico do dialecto crioulo do arquipélago de Cabo Verde*, Mindelo, 1991,

LANG Jürgen, «Estruturas eventualmente africanas no crioulo de Cabo Verde (Santiago)», in *Papia*, Brasília, 1994, Vol. III, n.º2, pp. 171-172.

PEREIRA, Dulce, *O universo do crioulo*, Setúbal, Escola Superior de Educação, 1993.

- PEREIRA, Dulce, "Textos de apoio à cadeira *Crioulos de Base Portuguesa*", Departamento de Linguística Geral e Românica da Faculdade de Letras da Universidade de Lisboa, edição policopiada, Lisboa, 1996.
- PIRES, João & HUTCHISON, John P., *Distonariu Preliminariu Kriolu. Preliminary Creole Dictionary Capeverdean/English*. Publikadu Pa Funk'Inhô Lobu, Boston, 1983.
- QUINT, Nicolas, *Dictionnaire français-cap-verdien*, Paris, L'Harmattan, 1997.
- QUINT [ABRIAL], Nicolas, *Lexique créole de Santiago-français*, Praia, Ed. de autor, 1996.
- QUINT, Nicolas, *Dictionnaire Cap-verdien -français*, Paris, L'Harmattan, 1999.
- SILVA, Baltasar Lopes da, *O dialecto crioulo de Cabo Verde*. Lisboa, Junta de Investigações do Ultramar, Centro de Estudos Políticos e Sociais, 1957.
- VEIGA, Manuel, *Diskrtson struktural di Língua kabuverdianu*. Praia, Instituto Kabuverdianu di Livru, 1982.
- VEIGA, Manuel, *Introdução à gramática do crioulo*, Praia, Instituto Caboverdiano do Livro e do Disco, 1995.
- Actas do Simpósio sobre a literatura e a cultura caboverdianas, Mindelo (Cabo Verde), 1986.

Bases de Dados Lexicais

- CORREIA, Margarita & GUERREIRO, Paula, "Bases de Dados Lexicais" in Mateus & Branco (orgs.) *Engenharia da Linguagem*, Edições Colibri, Lisboa, 1995.
- BARNBOOK, Geoff, *Language and Computers. A practical introduction to the computer analysis of language*, Edinburgh University Press, 1996.
- BOGURAEV, Bran & BRISCOE Ted, *Computational Lexicography for Natural Language Processing*, London and New York: Longman, 1989.
- FERNANDES, Carla, *Computational Lexicography: a contribution for new bilingual dictionaries" tese de mestrado apresentada à Universidade Nova de Lisboa, Faculdade de Ciências Sociais e Humanas, sob a orientação da Profª Doutora Maria Teresa R. F. Lino, Lisboa, 1995.*
- JACKSON, H. *Words and Their Meanings*, London, Longman, 1988.
- SINCLAIR, John (cd.) *An Account of the Cobuild Project in Lexical Computing*, London, Collins Cobuild, 1987.
- SINCLAIR, John "Introduction to the The Cobuild Dictionary" in *Collins Cobuild English Language Dictionary*, London, Collins Cobuild, 1987.

Dicionários utilizados

- Le Petit Robert, Dictionnaire de la Langue Française*, edição em CD-ROM. Liris Interactive.

ACTAS DO XV ENCONTRO NACIONAL DA APL

- MACHADO, José Pedro (coord.), *Dicionário da Língua Portuguesa*, Lisboa, Sociedade de Língua Portuguesa, 1964.
- FERREIRA, Aurélio Buarque de Holanda, *Novo Dicionário Aurélio da Língua Portuguesa*, 2ª edição revista e aumentada, Rio de Janeiro, Editora Nova Fronteira, 1986.
- Dicionário da Língua Portuguesa*, edição em CD-ROM, Porto Editora Multimedia e Priberam Informática, 1996.
- Lexlello*, 2ª edição revista e aumentada, Porto, Lello & Irmão - Editores, 1992.
- SANTOS, António Nogueira, *Novos Dicionários de Expressões Idiomáticas*, Lisboa, Edições João Sá da Costa, 1990.
- SIMÕES, Guilherme Augusto, *Dicionário de Expressões Populares Portuguesas*, Lisboa, Publicações D. Quixote, 1993.