

TRADUÇÃO AUTOMÁTICA NA INTERAÇÃO COM MÁQUINAS

ANTÓNIO TEIXEIRA, JOSÉ CASIMIRO PEREIRA,
PEDRO FRANCISCO E NUNO ALMEIDA

ABSTRACT

Automatic translation is usually related to conversion between human languages. Nevertheless, in human-machine interaction scenarios new forms of translation emerged. This work presents two examples. First, from the area of Natural Language Generation, is presented a data-to-text system, where data stored in a database regarding a medication plan is translated to Portuguese. As second example, is presented a system addressing the transmission of information from humans to computers, showing that automatic translation can be useful in the development of systems that use voice commands for interaction and having multilingualism as a requirement. The examples presented, part of our recent work, demonstrate the increase of application areas for automatic translation, area that received many and valuable contributions from Belinda Maia.

[1] INTRODUÇÃO

A tradução automática de linguagem é, em geral, associada à conversão entre línguas humanas. No entanto, sendo a interação com computadores (ou sistemas integrando estes, como os robôs), no essencial, “a transmissão de informação” e sendo as linguagens naturais a melhor forma, até hoje, criada pelo Homem para codificar informação - como defende, por exemplo, Santos (1992) “Natural language is so far the most comprehensive tool for (humans to) encode and reason with knowledge” - é natural que a tradução automática tenha papéis a desempenhar na nossa interação com as máquinas (e das máquinas connosco).

A nossa interação com as máquinas é, em geral, bidirecional. Tomemos como exemplo uma aplicação simples para nos informar sobre a previsão do tempo, a correr num dos cada vez mais omnipresentes *Smartphones* (que não nos atrevemos a traduzir...). A linguagem natural é passível de utilização na transmissão de informação da previsão para a semana, sob a forma de um texto ou mesmo pela leitura desse texto, usando um sintetizador de voz (dois exemplos do que se designa habitualmente como modalidades de saída). Outra utilização na interação consiste em navegar nas várias informações disponíveis utilizando comandos de voz, dizendo, por exemplo, “quero saber a previsão para os próximos dias” (exem-

plo de uma modalidade de entrada). Para tornar possíveis funcionalidades como as dos exemplos, torna-se necessário tecnologias capazes de:

- (i) traduzir o sinal acústico da voz em comandos com um significado preciso para a aplicação.
- (ii) traduzir (uma pequena provocação na utilização da palavra!) em frases, ou textos, informação que existe internamente nas máquinas;

[1.1] *Controlo por voz de aplicações multilingue*

A tecnologia atual da área do processamento de fala e de linguagem natural permite o controlo de aplicações e a procura de informação com um razoável desempenho. Permite também ditar mensagens, como o conteúdo de um email ou qualquer outro documento. Por forma a se conseguir utilizar a linguagem falada como modalidade/forma de entrada de informação ou controlo de uma aplicação de computador, é necessário efetuar pelo menos duas tarefas¹:

Primeiro, converter o sinal acústico numa sequência de palavras. Tarefa que é desempenhada pelos denominados reconhedores de fala. De seguida, a sequência de palavras é processada por um sistema de Compreensão de Fala (em inglês, *Spoken Language Understanding* – SLU) responsável por extrair informação semântica – significado.

Para que possam efetuar a sua tarefa, os reconhedores necessitam de informação sobre as palavras e frases que devem considerar no seu processo de reconhecimento. Em geral, para aplicações de controlo de aplicações, esta informação é fornecida por gramáticas.

Para a compreensão da linguagem foram propostos vários tipos de abordagens (ver [Tur & De Mori 2011](#)) com graus diferentes de complexidade. Um dos métodos mais utilizados baseia-se na utilização de gramáticas semânticas ([Tur & De Mori 2011](#), p. 51), sendo o sistema mais conhecido o Phoenix ([Ward 1990](#)).

Estas gramáticas são dependentes da língua e são criadas com base no contexto de uma aplicação, sendo necessário traduzir as gramáticas de cada aplicativo para cada idioma que a aplicação deve suportar. Atendendo a que, em geral, os programadores não possuem conhecimentos num conjunto alargado de línguas, e a tradução das gramáticas é algo custoso, mesmo por tradutores formados segundo a filosofia defendida pela Belinda de serem capazes de colaborar com “a máquina” ([Santos & Simões 2015](#)), torna-se pertinente explorar as capacidades da tradução automática nestes contextos, o que os autores têm vindo a fazer ([Teixeira et al. 2014a](#)).

[1] Pelo menos com a tecnologia mais comum. Existem propostas recentes de sistemas que integram estas duas partes.

[1.2] *Conversão de dados em frases*

Os sistemas de Geração de Linguagem Natural precisam de mapear alguma fonte de informação (como uma base de dados, por exemplo) em algum tipo de mensagem gerada automaticamente (Bateman & Zock 2003). Esse mapeamento pode ser feito recorrendo a diferentes tipos de sistemas, desde os baseados em modelos pré-definidos (*templates*) até aos sistemas baseados em aprendizagem automática. A escolha do tipo de sistema depende da aplicação, recursos e conhecimentos disponíveis. Para uma utilização em interação é essencial que as frases geradas apresentem variabilidade e que o desenvolvimento seja possível sem grandes conhecimentos de Linguística.

Se por um lado, a opção por modelos pré-definidos proporciona resultados rápidos, por outro verifica-se uma uniformidade nas respostas produzidas pelo sistema. Esta uniformidade torna este método desadequado para sistemas que pretendam um grau elevado de naturalidade na interação com o ser humano, como os que estamos interessados.

Os sistemas convencionais, não usando *templates*, profusamente descritos em (Reiter & Dale 2000), efetuam este mapeamento em 3 fases: (1) Planificação do Documento; (2) Planeamento Fino, incluindo Lexicalização (escolha das palavras que melhor descrevem o que foi definido na planificação do documento), Agregação e Geração de Expressões Relacionadas; (3) Realização de superfície, que concretiza todas as decisões tomadas nas fases anteriores.

Estes sistemas utilizam, frequentemente, regras bastante específicas, relacionadas com o domínio do problema que desejam solucionar, sendo fortemente dependentes da experiência e conhecimento do seu desenvolvedor. A sua adaptação a novos requisitos é, em geral, bastante difícil (Lemon 2010).

Tentando dar resposta a estas limitações, foram propostos, mais recentemente, sistemas usando aprendizagem automática. O objetivo é treinar/ensinar um sistema com base em informações previamente conhecidas. Depois, com este conhecimento adquirido, é possível inferir novas expressões. Expressões que, em geral, não fazem parte do conjunto de treino.

Diversas abordagens têm sido seguidas. Algumas utilizam ontologias como estrutura para recolher o corpus e posteriormente gerar as novas expressões (Stent et al. 2004; Stent & Molina 2009; Vogiatzis et al. 2008); outras formulam o processo de geração como um processo de tradução entre uma linguagem interna e uma língua natural (Langner & Black 2009; Langner 2010) e utilizam ferramentas de tradução automática como o MOSES (Koehn et al. 2007).

A geração de linguagem para conversão de dados (internos a um sistema) em texto (Data2Text em inglês), em que se pode incluir a geração para sistemas de diálogo, apresenta algumas diferenças importantes relativamente a outros sistemas de geração de linguagem natural (Langner 2010). Neste caso, a geração consiste quase exclusivamente na última fase, a realização de superfície.

As abordagens usando aprendizagem automática são particularmente interessantes para este cenário de geração “simplificado”.

Alguns exemplos de sistemas orientados à conversão de dados em texto:

- BabyTalk - Este sistema (Portet et al. 2009; Hunter et al. 2011) surgiu com o objetivo de apoiar os profissionais de saúde (enfermeiros e médicos) de uma Unidade de Cuidados Intensivos Neonatais, na sua necessidade de assimilar uma grande quantidade de informação, em muito pouco tempo, sobre os bebês aí internados.
- Mountain - Foi desenvolvido por Langner, como parte da sua tese de doutoramento e utiliza, igualmente, um corpus alinhado (Langner & Black 2009; Langner 2010). Utiliza o MOSES (Koehn et al. 2007; Koehn 2014) como forma de gerar as frases. A sua “linguagem de entrada” corresponde a uma sequência de códigos que representam a disponibilidade de um “court” de ténis. A “linguagem de saída” corresponde à tradução desse código em inglês.

000000 d5 t3	friday evening is completely closed
100000 d2 t2	the only time available is noon
111111 d4 t1	the court is open all morning
111111 d1 t3	you can reserve a court anytime on monday evening
100011 d5 t3	six , ten or eleven
010011 d3 t2	you can reserve a court at 1pm , 4pm and 5pm on wednesday
011001 d4 t3	any time but 6 , 9 and 10
111011 d7 d2	afternoon except the 3pm block
111100 d1 t2	you can reserve a court is free anytime from noon until 3
110111 d6 t3	saturday evening . ooh , that

FIGURA 1: Exemplo do corpus do Mountain (retirado de Langner (2010)).

- SINotas - Apresentado em (Araújo et al. 2010; Novais et al. 2009), é um dos poucos exemplos, de que tenhamos conhecimento, de sistemas deste género, orientados para uso da língua portuguesa. Desenvolvido no Brasil, destina-se a dar suporte a uma aplicação de divulgação de notas de alunos, numa universidade. O SINotas utiliza um corpus alinhado, onde a cada nota possível de um aluno associa uma descrição para essa nota.

[1.3] *O que se segue ...*

Neste artigo apresenta-se trabalho recente nestas duas aplicações das tecnologias da tradução automática - conversão de dados para texto e suporte à interação por voz multilingue - descrevendo as abordagens, trabalho recente e apresentando exemplos para o Português.

[2] TRADUÇÃO NA CONVERSÃO ENTRE DADOS E TEXTO PARA PORTUGUÊS

Os autores têm vindo a trabalhar num sistema de geração de linguagem natural para utilização em aplicações necessitando de transformar informação interna em frases passíveis de ser lidas pelo utilizador, ou mesmo convertidas em fala sintética através de um sintetizador. Este sistema, que se pretende seja de fácil aplicação a novos domínios, baseia-se na utilização de corpora paralelos para treino de um sistema de tradução.

O cenário escolhido para os primeiros testes a este sistema de geração foi o da informação sobre medicação a ser tomada (Ferreira et al. 2014; Pereira et al. 2012). Numa primeira fase, o problema foi simplificado, recorrendo apenas à geração de uma frase relativa a um único medicamento. Neste cenário, a informação relativa a cada um dos medicamentos a tomar, em determinada altura, pode ser visualizada sob a forma de uma frase, complementando outras formas de apresentação, como listas e imagens.

[2.1] Implementação

Em linhas gerais, o funcionamento do sistema é ilustrado na Figura 2):

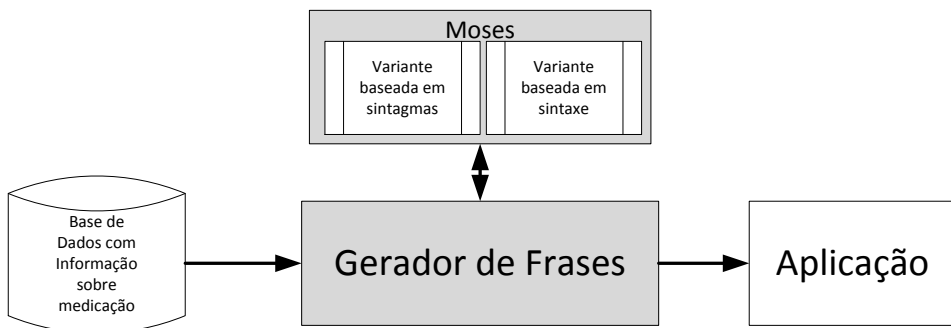


FIGURA 2: Funcionamento do sistema de conversão de informação de medicação em frases.

A parte central deste sistema é o módulo capaz de criar uma frase em resposta a um vetor com dados, fornecido como entrada. Para se alcançar este objetivo são necessários três componentes:

A base de dados com **informação relativa aos planos de medicação**, que designaremos por BD, é o componente responsável por armazenar informação sobre o plano(s) de medicação.

O componente **Moses** é responsável pela tradução para português das frases, enviadas pelo módulo gerador de frases. Para efetuar a tradução, o Moses tem de ser treinado com um corpus paralelo. A cada frase na “linguagem” de entrada deve corresponder uma frase na linguagem de saída, respeitando o ordenamento

dos ficheiros. A “linguagem” de entrada será constituída por valores correspondentes aos fornecidos pelo módulo BD. Na linguagem de saída estão as expressões, em português, que se deseja que o Moses seja capaz de gerar.

O módulo **gerador de frases** é responsável por receber os pedidos dos utilizadores e interagir com a base de dados. É também sua responsabilidade enviar mensagens escritas na linguagem de entrada para o módulo Moses e receber a resposta na linguagem de saída. Por último, compete-lhe processar as respostas e apresentá-las ao utilizador, através da respetiva Aplicação.

Como referido anteriormente, foi utilizado um corpus constituído por duas linguagens, alinhadas.

A “linguagem” de entrada reflete os dados que são obtidos por consulta, na base de dados, tendo sido selecionados 9 tipos de dados: Nome e Apelido do utilizador, Mensagem de cortesia, Nome do medicamento a tomar, Tipo do medicamento, Forma de tomar o medicamento, Cor do medicamento, Dose a tomar e Frequência da toma.

Correspondentemente, na linguagem de saída surge uma frase que exprime o mesmo tipo de informação da “linguagem” de entrada, mas em português. A Tabela 1 apresenta exemplos destas duas linguagens.

Linguagem interna	Frase correspondente
<p>peessoa32n saudacao_0 pessoa0a medicamento21 tipo0 tomar0 cor00 dose0 freq-toma00</p> <p>peessoa0n saudacao_0 pessoa0a medicamento14 tipo1 tomar2 cor00 dose4 freq-toma02</p> <p>peessoa0n saudacao_m pessoa12a medicamento0 tipo8 tomar3 cor00 dose0 freq-toma02</p> <p>peessoa78n saudacao_f pessoa0a medicamento3 tipo1 tomar2 cor00 dose3 freq-toma04</p>	<p>Helena pode tomar agora o Seretaide.</p> <p>Vai-se deitar então tome quatro comprimidos Primperan.</p> <p>Antes de deitar senhor Lima não se esqueça da bomba de inalação.</p> <p>Dona Teresinha está na hora de almoço tome os três comprimidos Ibuprofeno.</p>

TABELA 1: Excerto do corpus alinhado utilizado nas experiências.

Foram criadas duas versões do sistema, diferindo apenas no método de tradução: um baseado em sintagmas (*phrase-based*), outro adoptando a chamada tradução baseada em sintaxe (*syntax-based*). O corpus utilizado é constituído por 643 frases, obtidas pela expansão de 126 frases produzidas por voluntários. Os dados presentes na base de dados correspondem a medicamentos e prescrições reais.

[2.2] *Exemplos de resultados*

Nesta secção apresentam-se vários exemplos seleccionados de forma a ilustrar os vários tipos de resultados obtidos. Pretende-se familiarizar o leitor com o que de facto foi possível obter usando os dois tipos de sistemas.

Ilustrando as capacidades do sistema e a disparidade de qualidade das frases obtidas, começamos por apresentar exemplos de geração produzidos por diferentes entradas.

Usando como entrada o vetor

< pessoa45n saudacao0 ... medicamento17 ... cor00 dose0 freqtoma00 >
que se refere ao medicamento 17 (Pulmicort) e contém informação sobre dose, tipo, ..., obteve-se a não frase (1).

(1) Luís Pulmicort de tomar agora o.

Para outra entrada,

< pessoa61n ... medicamento4 tipo0 tomar0 cor00 dose2 freqtoma00 >
os resultados são muito mais aceitáveis (2).

(2) Senhor Paulo tome dois comprimidos de Maltofer.

Os resultados (1) e (2) e outros análogos podem ser comparados com os gerados por humanos, e que integram o corpus. Apresentam-se de seguida dois exemplos de saída do sistema (assinalados com S antes) alinhados com frases produzidas por humanos (com H antes). Estes exemplos mostram o alinhamento entre as frases criadas pelos humanos, e que servem de referência, e as geradas pelo sistema, para uma mesma entrada. As frases são, aqui, apresentadas em minúsculas para que seja possível evidenciar as suas diferenças, como a seguir explicado. Nos alinhamentos, a falha do sistema em incluir uma palavra na frase é marcada com “****”. Quando ocorrer troca de palavras, ou quando houver adição de palavras, estas diferenças são evidenciadas através de maiúsculas. Esta evidenciação é especialmente direccionada para os casos em que haja marcação com “****”.

(3) H: dona denise assim que se levantar não se esqueça de tomar OS COMPRIMIDOS nicotibine

S: dona denise assim que se levantar não se esqueça de tomar O COMPRIMIDO nicotibine

(4) H: DEVE TOMAR AGORA ao acordar **** a bomba de inalação DE pulmicort AUGUSTO

S:*** ** AUGUSTO ao acordar APLIQUE a bomba de inalação *** pulmicort ***

Em (3) a grande diferença resulta do sistema usar o singular para “OS COMPRIMIDOS”. Num cenário em que a frase gerada é a única informação que é trans-

mitida ao utilizador esta falha pode ter consequências muito graves ao potenciar a toma de uma dose errada. No entanto, em cenários multimodais em que a frase complementa outra informação clarificando o número de comprimidos a tomar, como é o caso dos que interessam aos autores, esta diferença pode não ter grandes consequências. No exemplo (4) existem várias diferenças, mas não se pode dizer que o sistema se tenha saído muito mal: “APLIQUE a bomba” será até mais adequado do que “TOMAR... a bomba”; “bomba de inalação pulmicort” tem a sua lógica. De referir que nesta fase optámos por não incluir vírgulas e outra pontuação na geração, devido a problemas com o analisador sintático (*parser*) utilizado. Sendo claramente uma limitação, foi já iniciada a exploração de *parsers* alternativos.

A diferença entre as frases geradas pelo sistema e as produzidas por humanos pode ser ainda maior. Os exemplos (5) e (6) mostram frases geradas inteligíveis e que se podem considerar boas em termos de naturalidade, mas que são completamente diferentes das produzidas por humanos. Este tipo de frases constitui um grande desafio para a avaliação, sendo normalmente consideradas como erros pelas métricas automáticas de avaliação, mas constituem, na opinião dos autores, uma das mais valias destes sistemas, afastando-os da monotonia dos sistemas de geração baseados em *templates*.

- (5) H: É HORA DE ALMOÇAR marcos não se esqueça de tomar *** quatro gotas de guttalex *** ***
S:*** *** *** *** marcos não se esqueça de tomar AS quatro gotas de guttalex AO ALMOÇO
- (6) H: *** *** É MEIO-DIA TOME AS três gotas de zaditen *** PATRÍCIA
S:PATRÍCIA NÃO SE ESQUEÇA DE TOMAR três gotas de zaditen AO MEIO-DIA

Em (7) e (8) exemplifica-se a diferença entre os resultados obtidos pelos 2 tipos de sistema. Utilizou-se a mesma informação de entrada na geração das frases. Notamos por “F” a geração baseada em sintagmas (*phrase-based*) e por “S” a geração baseada em sintaxe.

- (7) F: Patrícia não se esqueça de tomar três gotas de Zaditen ao meio-dia
S: Patrícia ao Zaditen gotas tome de três meio-dia
- (8) F: Senhora Carvalho após o seu almoço tome cinco comprimidos de Duphaston
S: Senhora Carvalho Duphaston comprimido branco cinco almoço

Destes exemplos transparece uma pior capacidade da variante do sistema baseada em sintaxe. Possivelmente, pelo tamanho bastante limitado do corpus utilizado e pelo facto de não se ter ainda conseguido um bom desempenho da anotação sintática. Estes resultados estão a ser encarados pelos autores não como uma

prova de que esta variante do sistema tem menor potencial, mas como um desafio para melhorar o desempenho dos processos adicionais que envolve.

[3] TRADUÇÃO NO SUPORTE À INTERAÇÃO POR VOZ MULTILINGUE

Como referido na introdução, uma forma de configurar o conhecimento e compreensão de fala é através de gramáticas, definidas para a aplicação em vista. Adoptamos essa abordagem para o desenvolvimento de diversas aplicações suportando a interação por voz, destacando-se o assistente AALFred (Saldanha et al. 2013; Teixeira et al. 2014b) do projeto AAL PaeLife. Para que a interação possa ser efetuada em múltiplas linguagens - o AALFred suporta atualmente Inglês, Português, Francês, Húngaro e Polaco - é definida uma gramática semântica de base e as gramáticas para as outras línguas são obtidas por tradução, seguida de verificação manual durante o desenvolvimento. As gramáticas necessárias para o reconhecedor de fala são também derivadas automaticamente.

[3.1] Implementação

Por forma a dar resposta às necessidades de desenvolvimento distribuído de aplicações e facilitar a verificação manual, o sistema foi implementado como um serviço web (*webservice*) e um portal web associado.

O sistema (Figura 3) é dual na funcionalidade. Suporta o desenvolvimento e o uso em contextos de interação real, altura em que as gramáticas nas várias línguas ficam disponíveis para utilização, sendo selecionadas em função da língua em utilização.

Em contextos de interação, o sistema é responsável pela compreensão da linguagem natural, aproveitando as gramáticas enviadas para o serviço em fase de desenvolvimento. Recebe a saída de reconhecimento de fala e retorna as informações semânticas extraídas. Também retorna, a pedido, as informações necessárias sobre as palavras e frases necessárias para configurar o reconhecedor de fala.

Dadas as limitações das traduções automáticas, o serviço também oferece suporte a revisão manual e atualização subsequente de gramáticas. Esta utilização é particularmente adequada quando se está na fase de desenvolvimento de uma aplicação, como o AALFred, ao permitir que cada parceiro envolvido no projeto possa rever e corrigir as gramáticas geradas automaticamente.

Todas as operações são feitas através do acesso a APIs², garantindo um controlo de operação consistente e completo.

Para permitir a introdução de novas gramáticas, uma interface específica é necessária para o desenvolvedor. Esta interface permite submeter uma gramática e verificar os resultados da sua tradução, tanto em termos de gramática gerada como de frases geradas por ela.

[2] Application Programming Interface (em português: Interface de Programação de Aplicações)

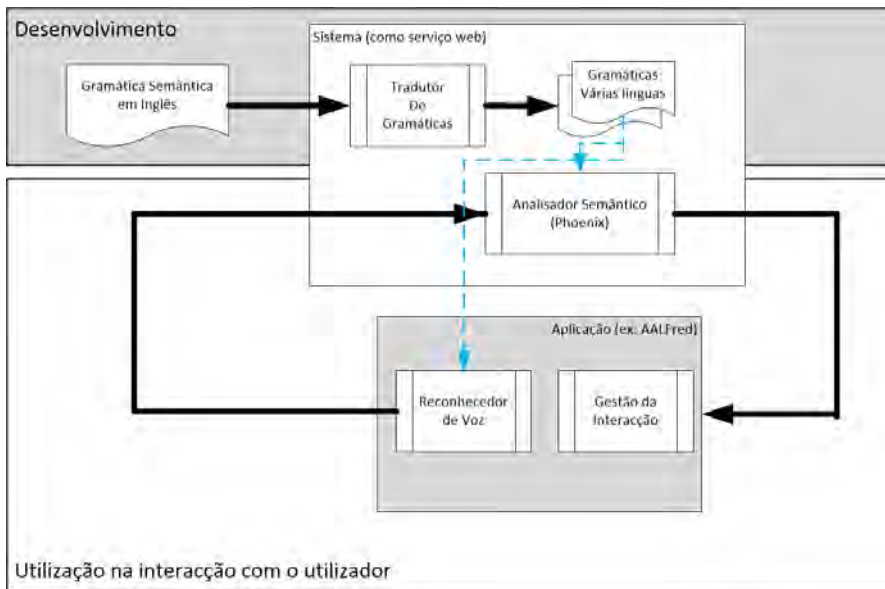


FIGURA 3: Sistema de análise semântica multilingue, baseado na tradução de gramáticas semânticas.

[3.2] Gramáticas semânticas e parser

O Phoenix (Ward 1990) foi escolhido como o analisador (*parser*) tendo sido também adoptado o seu formato de especificação de gramáticas. A escolha teve por base a robustez do Phoenix a erros no reconhecimento e desempenho e versatilidade que demonstrou em variadíssimas aplicações.

O sistema de análise semântica Phoenix (Ward 1990) modela diretamente a semântica de um domínio específico usando gramáticas semânticas baseadas em quadros (*frames*) e *slots*. Cada *slot* tem uma gramática livre de contexto associado, que especifica padrões de sequências de palavras que coincidem com o *slot* e é compilada como uma rede de transição recursiva (RTN). São preenchidos através da comparação entre a sequência de palavras das frases em análise com estas redes recursivas (Tur & De Mori 2011, p. 51).

O objetivo do analisador (*parser*) é extrair as anotações semânticas (*tags*), conforme definido na gramática semântica. Esta operação é efetuada sobre a lista de palavras que foi fornecida pelo sistema de reconhecimento de fala. Após esta tarefa, o texto juntamente com as respetivas anotações é enviado para processamento pelo Gestor de Interação. Por último, o resultado final é usado pela aplicação.

As gramáticas do Phoenix contêm as regras livres de contexto que especificam os padrões da palavra. Uma pequena gramática exemplo é apresentada em (9).

Trata-se de um exemplo muito simples que servirá para a explicação do processo de tradução — apresentado na próxima subsecção — que permite transformar sequências de palavras como “show this contact’s pictures” numa interpretação do que o utilizador pretende: que lhe seja mostrada a foto do contacto atualmente selecionado:

```
(9)  [Main]
      ([AGENDA])
      ([CONTACTS])
      ;
      [CONTACTS]
      (show this contact's [PHOTOS])
      ;
      [PHOTOS]
      (photographies)
      (photos)
      (pictures)
      ;
```

As regras, uma por linha, aparecem entre parêntesis curvos. Nomes entre parêntesis rectos indicam não-terminais. Palavras em minúsculas indicam símbolos terminais. É possível indicar que algo é opcional utilizando o “*” ou que pode ter uma ou mais ocorrências usando o “+”.

[3.3] *Processo de tradução das gramáticas*

O principal objetivo deste processo é gerar automaticamente uma gramática derivada em outros idiomas de destino. Isso é conseguido através da preservação, tanto quanto possível, da estrutura da gramática, gerando frases coerentes na língua-alvo e tendo em consideração o processo de reordenação de palavras.

O objetivo é traduzir para uma língua-alvo todas as palavras terminais preservando as anotações (*tags*) semânticas. A tradução também deve produzir uma lista completa das frases definidas pela gramática.

O processo adotado e implementado é composto por três etapas: 1) expansão completa da gramática; 2) tradução; e 3) reconstrução de gramática.

Expansão da gramática

Com o objetivo de manipular as gramáticas do Phoenix, uma de duas abordagens teria de ser seguida: alterar o Phoenix ou ter um analisador separado. Foi decidido implementar um analisador separado e contar com o analisador de Phoenix apenas para a sua função já especificada e bem testada: análise de texto de entrada com base em uma gramática definida.

Para traduzir corretamente a gramática, é necessário apresentar frases completas ao tradutor, para assim ter forma de avaliar adequadamente qual a tradução mais plausível. Para obter essas frases, o algoritmo desenvolvido faz uso de duas estruturas de dados: uma pilha “em andamento” e uma fila “feito até agora”. Na primeira, o algoritmo armazena a regra atual, enquanto no segundo armazena as palavras traduzidas. A expansão de todas as regras é feita mantendo o histórico das regras visitadas ao longo da expansão.

O nosso exemplo será expandido como em (10).

- (10) show this contact's photographs
show this contact's photos
show this contact's pictures

Tradução

O processo de tradução consiste em submeter o resultado da expansão (palavras mais as suas regras gramaticais/história) e receber as frases traduzidas resultantes (emparelhamento de palavras na tradução com as palavras correspondentes na fonte).

Para a tradução, a escolha recaiu sobre o tradutor Bing (Microsoft 2014), utilizado através da Microsoft Translator API (Microsoft 2015), devido à sua capacidade de fornecer informação sobre a reordenação das palavras. Esta informação facilita a correspondência das palavras da tradução com palavras de origem, essencial para a reordenação das palavras aquando da reconstrução das regras. Além disso, este tradutor também permite obter múltiplas traduções por pedido, o que permite a expansão de uma gramática existente para oferecer suporte a várias frases semelhantes, sem a necessidade de entrada adicional. Podemos, assim, aumentar a cobertura da nossa gramática de forma automática e sem esforço.

Em (11) apresentam-se as frases resultantes da tradução da expansão apresentada anteriormente, em (10).

- (11) mostrar fotos deste contacto
mostrar fotografias deste contacto

Reconstrução da gramática

Quando a gramática é analisada (a fim de expandi-la depois), um objeto diferente é criado para cada instância de qualquer regra. Como tal, para cada palavra terminal presente na instrução resultante da expansão da gramática, podemos determinar exatamente qual a regra que deu origem ao caminho que leva a ela, após a tradução. Como temos informação relativa à reordenação disponível, sabemos quais as regras que geraram o texto resultante da tradução.

O algoritmo desenvolvido utiliza a história de expansão da gramática e as frases traduzidas. Consiste em analisar informações de histórico dos antepassados para refazer a gramática. Isto é feito através da fusão de não-terminais, do mesmo nível, em toda a gramática numa abordagem de cima para baixo. As Figuras 4 e 5 ilustram as fases inicial e final deste processo.

[<i>Main</i>] [<i>CONTACTS</i>] mostrar	[<i>Main</i>] [<i>CONTACTS</i>] [<i>PHOTOS</i>] fotos	[<i>Main</i>] [<i>CONTACTS</i>] deste	[<i>Main</i>] [<i>CONTACTS</i>] contacto
---	--	---	--

FIGURA 4: Representação da informação na fase inicial do processo de reconstrução da gramática.

[<i>Main</i>] [<i>CONTACTS</i>] mostrar	[<i>PHOTOS</i>] fotos	deste	contacto
---	----------------------------	-------	----------

FIGURA 5: Representação da informação na fase final do processo de reconstrução da gramática.

[3.4] *Exemplo de utilização*

Este sistema foi adoptado pelo assistente pessoal AALFred (Saldanha et al. 2013; Teixeira et al. 2014b) desenvolvido no projeto AAL PaeLife. O AALFred é uma aplicação com interação multimodal (voz, toque, gestos, teclado, rato) que facilita a utilização das redes sociais, das múltiplas formas de comunicação (email, twitter, Skype...) e acesso a informação (meteorologia, notícias, ...) assim como a gestão da agenda e contactos.

As gramáticas semânticas e as suas traduções atrás descritas suportam toda a interação por voz do AALFred. A título de exemplo, apresenta-se, a seguir, a subgramática referente a funcionalidades da agenda, na versão inicial em inglês (12) e o resultado do processo de tradução atrás descrito em (13).

- (12) [AGENDA]
- (agenda)
- ([CHANGEDATE])
- (go to my agenda)
- (*go *to [NEXT] [DATEELEMENT])
- (*go *to [PREVIOUS] [DATEELEMENT])

```

(i want to see my agenda)
(*open [WEEKDAYS])
(show my agenda)
;
(13) [AGENDA]
(abre a minha agenda)
(abre [WEEKDAYS])
(abrir *a agenda)
(abrir a minha agenda)
(*abrir [WEEKDAYS])
(agenda)
([CHANGEDATE])
(*eu quero ver a minha agenda)
(ir para a minha agenda)
(ir para *a [NEXT] [DATEELEMENT])
(ir para *a [PREVIOUS] [DATEELEMENT])
(ir para [WEEKDAYS])
(mostra-me a minha agenda)
(mostra-me [WEEKDAYS])
(mostra a minha agenda)
(mostrar a minha agenda)
(mostrar [WEEKDAYS])
([NEXT] [DATEELEMENT])
([PREVIOUS] [DATEELEMENT])
(quero ver [WEEKDAYS])
;

```

[4] CONCLUSÕES

Neste artigo apresentam-se dois exemplos de aplicação de tradução automática na área da interação entre humanos e máquinas em que existe trabalho recente para o português.

O primeiro exemplo, na área da geração de linguagem, faz uso de um sistema de tradução automática, o MOSES, para a conversão de informação relativa a planos de medicação em frases em português.

O segundo, do lado oposto da interação - transmissão de informação entre humanos e máquinas por oposição à transmissão entre as máquinas e os humanos do primeiro exemplo -, mostra quão útil pode ser a tradução no desenvolvimento de aplicações que incluam comando por voz e que tenham por objetivo suportar múltiplas línguas. Atendendo às limitações atuais dos sistemas de reconhecimento e de compreensão, a tarefa de tradução é ainda simples. No entanto, com o evoluir

destas tecnologias e a adoção de sistemas capazes de dialogar com o utilizador, terão de ser contempladas as diferenças entre as línguas e culturas. Por exemplo³, em algumas línguas há mais preliminares o que deverá implicar dimensões diferentes para cada bloco ou mesmo a necessidade de adicionar blocos inexistentes na língua original. Consideramos que o protótipo existente pode desempenhar um papel relevante na criação de corpos comparáveis, ao permitir recolha das interações em situações similares para diferentes línguas.

Estes dois exemplos mostram a utilidade crescente dos sistemas de tradução automática, mesmo para a nossa comunicação com as máquinas. Estas possibilidades só se tornaram possíveis com o trabalho de muitos, em que se inclui e destaca a Belinda, para que a tradução, automática ou não, evoluísse.

Para terminar, consideramos que a relação entre máquinas, humanos e tradução vem acrescentar ainda mais riqueza à relação que Belinda Maia sempre tem considerado ser benéfica entre tradutores (humanos) e máquinas, em que as máquinas podem ajudar humanos na tradução. Nos exemplos apresentados, a tradução ajuda à comunicação/interação entre esses mesmos humanos e as mesmas, ou outras, máquinas.

AGRADECIMENTOS

Os autores agradecem a todos os que contribuíram para a criação do corpus e a todos os que participaram na avaliação das frases que tornaram possível o trabalho na geração de frases. Um agradecimento especial ao Mário Rodrigues pela ajuda na obtenção e utilização do analisador sintático para o português.

Relativamente ao trabalho na tradução das gramáticas semânticas, os autores não podem deixar de agradecer a todos os parceiros do projeto AAL PaeLife, e em especial ao Microsoft Language Development Center (MLDC), pela ajuda na definição de requisitos, pelo retorno que nos forneceram, e pela adoção deste componente no AALFred.

Os autores agradecem a preciosa ajuda na revisão do texto de Samuel Silva.

O trabalho mencionado neste artigo foi parcialmente financiado pelo FEDER, COMPETE and FCT através dos projetos AAL/0015/2009, AAL PaeLife, QREN AAL4ALL e financiamento à unidade de investigação IEETA (PEst-OE/EEI/UI0127/2014).

Os nossos agradecimentos, também, aos Editores deste volume pelo convite - que muito nos honra -, pela ajuda, comentários, disponibilidade em todo o processo e, muito mais importante, por se dedicarem a esta nobre iniciativa.

[3] Agradecemos à Diana Santos pela ideia.

REFERÊNCIAS

- Araújo, Roberto, Rafael Oliveira, Eder Novais, Thiago Tadeu, Daniel Pereira & Ivandré Paraboni. 2010. SINotas: the Evaluation of a NLG Application. Em *Proceedings of the Seventh International Conference on Language Resources and Evaluation (LREC)*, 2388–2391.
- Bateman, John & Michael Zock. 2003. Natural language generation. Em Ruslan Mitkov (ed.), *The Oxford Handbook of Computational Linguistics*, 284–304. Oxford University Press.
- Ferreira, Flávio, Nuno Almeida, Ana Filipa Rosa, André Oliveira, José Casimiro Pereira, Samuel Silva & António Teixeira. 2014. Elderly centered design for interaction - the case of the S4S medication assistant. Em *Procedia Computer Science*, vol. 27, 398–408.
- Hunter, James, Yvonne Freer, Albert Gatt, Ehud Reiter, Somayajulu Sripada, Cindy Sykes & Dave Westwater. 2011. BT-Nurse: computer generation of natural language shift summaries from complex heterogeneous medical data. *Journal of the American Medical Informatics Association (JAMIA)* 18. 621–624.
- Koehn, Philipp. 2014. *MOSES: Statistical Machine Translation System - User Manual and Code Guide*. <http://www.statmt.org/moses/manual/manual.pdf>.
- Koehn, Philipp, Hieu Hoang, Alexandra Birch, Chris Callison-Burch, Marcello Federico, Wade Shen, Christine Moran, Richard Zens, Ondřej Bojar, Alexandra Constantin & Evan Herbst. 2007. Moses: Open source toolkit for statistical machine translation. Em *45th annual meeting of the association for computational linguistics (demo and poster sessions)*, 177–180.
- Langner, Brian. 2010. *Data-driven Natural Language Generation: Making Machines Talk Like Humans Using Natural Corpora*: Carnegie Mellon University. Tese de Doutoramento.
- Langner, Brian & Alan W. Black. 2009. MOUNTAIN: A Translation-based Approach to Natural Language Generation for Dialog Systems. Em *First International Workshop on Spoken Dialogue Systems Technology (IWSDS)*, s/pp.
- Lemon, Oliver. 2010. Learning what to say and how to say it: joint optimization of spoken dialogue management and natural language generation. *Computer Speech & Language* 25. 210–221.
- Microsoft. 2014. Bing translator. <http://www.bing.com/translator/>.
- Microsoft. 2015. Microsoft translator API. <http://www.microsoft.com/translator/translator-api.aspx>.

- Novais, Eder, Rafael Oliveira, Daniel Pereira & Thiago Tadeu. 2009. A Testbed for Portuguese Natural Language Generation. Em *Seventh Brazilian Symposium in Information and Human Language Technology*, 154 – 157.
- Pereira, José Casimiro, António Teixeira & Joaquim Sousa Pinto. 2012. Natural Language Generation in the context of Multimodal Interaction in Portuguese. *Electrónica e Telecomunicações* 5. 400–409.
- Portet, François, Ehud Reiter, Alberto Gatt, Jim Hunter, Somayajulu Sripada, Yvonne Freer & Cindy Sykes. 2009. Automatic generation of textual summaries from neonatal intensive care data. *Artificial Intelligence* 173. 789–816.
- Reiter, Ehud & Robert Dale. 2000. *Building natural language generation systems*. Cambridge University Press.
- Saldanha, Nuno, Jairo Avelar, Miguel Dias, António Teixeira, Daniel Gonçalves, Emmanuel Bonnet, Karine Lan, Németh Géza, Petra Csobanka & Artur Koleinski. 2013. A Personal Life Assistant for “natural” interaction: the PaeLife project. Em *AAL Forum*, poster presentation.
- Santos, Diana. 1992. Natural Language and Knowledge Representation. Em *Proceedings of the ERCIM Workshop on Theoretical and Experimental Aspects of Knowledge Representation*, 195–197.
- Santos, Diana & Alberto Simões. 2015. Ensinador paralelo: Alicerces para uma pedagogia nova. Neste volume.
- Stent, Amanda & Martin Molina. 2009. Evaluating automatic extraction of rules for sentence plan construction. Em *Proceedings of the SIGDIAL 2009 Conference: 10th Annual Meeting of the Special Interest Group on Discourse and Dialogue*, 290–297.
- Stent, Amanda, Rashmi Prasad & Marilyn Walker. 2004. Trainable sentence planning for complex information presentation in spoken dialog systems. Em *Proceedings of the 42nd annual meeting on association for computational linguistics*, 79–86.
- Teixeira, António, Pedro Francisco, Nuno Almeida, Carlos Pereira & Samuel Silva. 2014a. Services to support use and development of speech input for multilingual multimodal applications for mobile scenarios. Em *The Ninth International Conference on Internet and Web Applications and Services (ICIW), Track Web Services-based Systems and Applications*, 41–46.
- Teixeira, António, Annika Hämmäläinen, Jairo Avelar, Nuno Almeida, Géza Németh, Tibor Fegyó, Csaba Zainkó, Tamás Csapó, Bálint Tóth, André Oliveira & Miguel Sales Dias. 2014b. Speech-centric multimodal interaction for easy-to-access online services. Em *Procedia computer science*, vol. 27, 389–397.

Tur, Gokhan & Renato De Mori. 2011. *Spoken language understanding: Systems for extracting semantic information from speech*. John Wiley & Sons.

Vogiatzis, Dimitrios, Dimitrios Galanis, Vangelis Karkaletsis, Ion Androutsopoulos & C. D. Spyropoulos. 2008. A Conversant Robotic Guide to Art Collections. Em *Proceedings of the 2nd workshop on Language Technology for Cultural Heritage Data (LREC)*, s/pp.

Ward, Wayne. 1990. The CMU air travel information service: Understanding spontaneous speech. Em *Proceedings of the darpa speech and natural language workshop*, 127–129.

CONTACTOS

António Teixeira

Departamento de Electrónica Telecomunicações e Informática/IEETA

Universidade de Aveiro

ajst@ua.pt

José Casimiro Pereira

Instituto Politécnico de Tomar

casimiro@ipt.pt

Pedro Goucha Francisco

IEETA, Universidade de Aveiro

goucha@ua.pt

Nuno Almeida

Departamento de Electrónica Telecomunicações e Informática/IEETA

Universidade de Aveiro

nunoalmeida@ua.pt