

Um Convite ao Processamento de Línguas de Sinais*

Antônio Carlos da Rocha Costa¹²³, Marianne Rossi Stumpf²
Juliano Baldez de Freitas¹, Graçaliz Pereira Dimuro¹

¹Escola de Informática – UCPel, RS, Brasil.

²PGIE/UFRGS, RS, Brasil.

³PPGC/UFRGS, RS, Brasil.

{rocha,liz,jubafreitas}@atlas.ucpel.tche.br

stumpf@qbnet.com.br

Abstract. *Este artigo faz um convite à comunidade de Processamento de Línguas Naturais (PLN) para que dirija o olhar ao tema do Processamento de Línguas de Sinais (PLS). As línguas de sinais são as línguas de modalidade gestual-visual utilizadas pelos surdos em sua comunicação diária. Há toda uma área de pesquisa dentro da Linguística, iniciada na década de 1960, voltada para o estudo das línguas de sinais. Com a utilização do sistema de escrita de línguas de sinais chamado SignWriting, é possível transpor imediatamente as técnicas de PLN e de Linguística Computacional voltadas para a análise e síntese de textos de línguas orais para a análise e síntese de textos escritos em línguas de sinais. O artigo resume inicialmente as características lingüísticas das línguas de sinais, apresenta a proposta de abordagem ao PLS fundamentada na utilização do sistema SignWriting, faz um levantamento dos desenvolvimentos de software para o sistema SignWriting que estão em andamento e, finalmente, apresenta alguns resultados iniciais da abordagem proposta, exemplificando a transposição de técnicas elementares de PLN para o domínio do PLS.*

1. Introdução

Este artigo fornece os elementos necessários para a compreensão da proposta de aplicação imediata das técnicas de Processamento de Línguas Naturais (PLN) ao domínio das línguas de sinais utilizadas pelos surdos na sua comunicação diária. Chamamos de Processamento de Línguas de Sinais (PLS) o resultado dessa transposição.

Os dois principais fundamentos dessa proposta de abordagem ao PLS são o sistema SignWriting [Sutton 1999] de escrita de línguas de sinais e o formato SWML [Costa 2004], um padrão baseado em XML que suporta adequadamente o armazenamento, recuperação e processamento de textos de línguas de sinais escritos no sistema SignWriting.

O artigo é essencialmente um convite à comunidade brasileira de PLN para que lance um olhar na direção das línguas de sinais (especialmente sobre a LIBRAS, a Língua Brasileira de Sinais), descobrindo a riqueza lingüística das mesmas e a oportunidade de pesquisa e desenvolvimento de aplicações de PLN que podem ser extremamente úteis do ponto de vista social e educacional, para os surdos. De certa forma, o artigo é também um balanço do Projeto SignNet, em que os autores vem trabalhando a questão do PLS desde 1996, com diversos apoios financeiros institucionais (universidades, agências de fomento, etc.). Esse Projeto deu início à divulgação do sistema SignWriting no Brasil e fomentou o desenvolvimento da proposta de PLS que está sendo resumida aqui.

A Seção 2 traz uma breve caracterização lingüística das línguas de sinais, enfatizando o seu caráter de línguas naturais e, portanto, a aplicabilidade que se pode fazer a elas das técnicas de PLN e de Linguística Computacional já existentes. A Seção 3 apresenta as características gerais do sistema SignWriting de escrita de línguas de sinais, colocando-o no contexto dos demais sistemas que foram propostos para finalidades semelhantes. A Seção 4 faz um levantamento dos recursos disponíveis atualmente para suporte ao PLS, destacando a importância internacional das contribuições para o PLS que tem sido produzidas no Brasil. A Seção 5 ilustra, através de um exemplo elementar de PLS, a proposta de transposição para as línguas de sinais das técnicas de PLN. O objetivo do exemplo elementar não é o de apresentar um resultado significativo na área de PLS, mas sim o de apresentar um caso simples, facilmente compreensível, de transposição para o domínio das línguas de sinais de um conjunto de técnicas de PLN desenvolvidas para o domínio das línguas orais. A Seção 6 traz a Conclusão do artigo.

*Trabalho parcialmente financiado pela FAPERGS e CNPq.



Figura 1: A expressão LÍNGUA-SINAIS-BRASIL escrita no sistema SignWriting.

2. A Caracterização Lingüística das Línguas de Sinais

William C. Stokoe [Stoke et. al. 1965, Maher 1996] foi o primeiro lingüista a realizar um estudo sistemático das línguas de sinais. Como tema de estudo, tomou a Língua de Sinais Americana (ASL). Mostrou sua estruturação lingüística nos diversos níveis (fonético, fonológico, morfológico e léxico) e deu indicações sobre sua estruturação sintagmática.

No nível fonético, as línguas de sinais se caracterizam por adotarem um aparelho articulador constituído pelas mãos, braços, rosto, tronco e pernas que realizam os gestos, expressões faciais e expressões corporais que constituem os sinais de uma língua de sinais.

No nível fonológico, os fonemas constituintes dos sinais são caracterizados principalmente pelos seguintes parâmetros fonológicos: as configurações de mãos, a orientação das mãos, os movimentos das mãos e dos dedos, a localização espacial onde esses movimentos são realizados, os pontos de contato entre as mãos e as outras partes do corpo, e as expressões faciais.

Os sinais são conjuntos organizados de fonemas. Cada sinal se constitui de uma sucessão de fonemas, onde se alternam fonemas individualizados e conjuntos de fonemas que tem ocorrência simultânea.

A caracterização léxica das línguas de sinais dá origem aos dicionários de línguas de sinais, que procuram descrever cada sinal individualmente e dar seu significado em termos de expressões de uma língua oral.

O primeiro desses dicionários foi o elaborado pelo próprio Stokoe. Trazia como novidade importante a criação de uma notação para a representação simbólica sistemática dos sinais, de um modo análogo ao modo como o alfabeto possibilita a escrita das palavras das línguas orais.

A caracterização sintagmática das línguas de sinais é área ainda a ser muito explorada, havendo bastante debate sobre questões, como por exemplo, a da ordem preferencial das partes do discurso nas frases de cada língua de sinais.

Dados adicionais sobre a caracterização lingüística das línguas de sinais podem ser encontrados no livro texto de Valli & Lucas [Valli & Lucas 1995].

3. O Sistema de Escrita SignWriting

Há diversos sistemas de escrita para línguas de sinais em utilização na atualidade. Além do sistema de Stokoe [Stoke et. al. 1965], que foi o primeiro a ser criado, como já se disse acima, outros sistemas amplamente utilizados são o HamNoSys [Hanke 2000] e o SignWriting [Sutton 1999, Sutton 2004]. Em nosso trabalho [Costa & Dimuro 2001, Costa & Dimuro 2002, Costa & Dimuro 2003, Costa, Dimuro & Freitas 2004], adotamos esse último para a escrita de línguas de sinais.

O SignWriting é um sistema de escrita com características gráfico-esquemáticas, que permite uma representação de textos de línguas de sinais através de uma forma intuitiva e de fácil compreensão. O sistema é constituído de um conjunto de símbolos e um conjunto de regras de escrita definidos para representar os diversos aspectos fonético-fonológicos das línguas de sinais. Desse modo, o SignWriting apresenta a feição de um sistema de escrita fonética para línguas de sinais, mas plenamente apto a suportar a delimitação de um subsistema de escrita de línguas de sinais que tenha características estritamente fonológicas.

A Figura 1 ilustra a escrita de uma expressão da Língua Brasileira de Sinais (LIBRAS), usando o sistema SignWriting. A expressão é LÍNGUA-SINAIS-BRASIL, que significa, justamente, “Língua Brasileira de Sinais”.

4. Recursos Existentes para Suporte ao PLS

O desenvolvimento de software para o PLS baseado em SignWriting é um fato recente. À exceção do editor de textos SignWriter [Sutton & Gleaves 1995], iniciativa pioneira datada do período 1985–1995, os demais programas e sistemas iniciaram seus desenvolvimentos somente após o ano 2000. Exatamente por isso, o SignWriter é ainda o programa mais utilizado pela comunidade de escrita de línguas de sinais [Sutton 2004].

Classificamos os recursos computacionais existentes para suporte ao PLS baseado em SignWriting em oito categorias:

- recursos para interoperabilidade (SWML e Unicode)
- editores de textos
- bancos de dados de sinais (dicionários)
- sintetizadores de animações de sinais
- reconhecedores automáticos de símbolos do sistema SignWriting
- tradutores de línguas de sinais baseados em glosas
- programas para comunicação online em línguas de sinais
- programas para processamento de textos escritos em línguas de sinais

4.1. Suporte à Interoperabilidade

As duas propostas existentes de suporte à interoperabilidade de sistemas baseados no SignWriting são a codificação dos símbolos do sistema SignWriting em Unicode e a utilização do padrão SWML [Costa 2004] para representação de textos de línguas de sinais em arquivos com formato baseado em XML.

O esforço de codificação Unicode para o SWML foi precedido pelo registro junto à ISO dos códigos para as línguas de sinais de todo o mundo, baseado nos códigos para as línguas orais dos países correspondentes (p.ex., a Língua Brasileira de Sinais é codificada por sgn-BR). A norma que especifica essa codificação é a ISO 639-2. O desenvolvimento da codificação dos símbolos propriamente dita, porém, não avançou significativamente deste então.

O formato SWML teve sua definição preliminar estabelecida em 2002 [Costa 2004]. Desde então ele passou a ser considerado e utilizado em diversas iniciativas de desenvolvimento de programas para processamento de textos e dicionários de línguas de sinais, como será visto a seguir.

A Figura 2 mostra a codificação SWML do texto da Figura 1. Note-se a representação individualizada de cada sinal, dentro dos elementos `<signbox>`, assim como a representação das diversas características dos símbolos que compõem os sinais, dentro dos elementos `<symbol>`. Note-se, igualmente, a presença das caixas de texto (`<textbox>`), contendo textos em língua oral, indicando um dos recursos de edição complementares, necessários a um editor de textos de línguas de sinais.

4.2. Editores de Texto

O principal editor de textos de línguas de sinais em uso hoje em dia ainda é o pioneiro dos programas para escrita de línguas de sinais, o SignWriter [Sutton & Gleaves 1995]. Trata-se de um programa para a plataforma Microsoft-DOS, com interface gráfica controlada pelo teclado, através de combinações de teclas.

O design do “teclado virtual” que permite a escrita dos textos é excelente, o que faz com que o processo de escrita se faça através de uma verdadeira digitação. Essa característica terminou gerando um público usuário com características bem marcadas de preferência por editores com interfaces baseadas em teclado.

Em conseqüência, a nova versão do editor, chamada de SignWriter Java [Sutton 2004], por ser implementada na linguagem Java, mantém essa preferência pela entrada de textos através da digitação.

O SW-Edit [Torchelsen & Costa 2002, Torchelsen & Costa 2004] é um programa multiplataforma, baseado em interface gráfica, controlado por mouse, que faz uso de todos os recursos disponíveis atualmente para as GUIs e que transporta todas as funcionalidades de edição dos modernos editores de texto para o contexto da escrita de línguas de sinais.

Tanto o SW-Edit como o SignWriter Java seguem a tradição criada pelo SignWriter, de fornecer ao usuário recursos para a criação de bancos de sinais (“dicionários”) que agilizam a escrita dos textos pela utilização de sinais previamente armazenados nesses dicionários. O SW-Edit faz uso exclusivo do formato SWML para armazenamento dos arquivos de textos e dos dicionários. O SignWriter Java vai utilizar SWML para exportação/importação de textos e dicionários gerados por outros programas.

```

<page height="1200" width="750">
  <line x="0" y="0" height="150" width="750">
    <cell x="150" y="0" height="150" width="150" bgcolor="255,255,255" cornersColor="0,0,0">
      <signbox>
        <symb x="45" y="59" x-flop="0" y-flop="1" color="0,0,0">
          <category>01</category>
          <group>05</group>
          <symbnum>001</symbnum>
          <variation>01</variation>
          <fill>02</fill>
          <rotation>01</rotation>
        </symb>
        <symb x="79" y="33" x-flop="1" y-flop="1" color="0,0,0">
          <category>02</category>
          <group>10</group>
          <symbnum>003</symbnum>
          <variation>01</variation>
          <fill>01</fill>
          <rotation>01</rotation>
        </symb>
        <symb x="46" y="19" x-flop="0" y-flop="1" color="0,0,0">
          <category>02</category>
          <group>10</group>
          <symbnum>005</symbnum>
          <variation>01</variation>
          <fill>02</fill>
          <rotation>01</rotation>
        </symb>
        <symb x="67" y="66" x-flop="0" y-flop="0" color="0,0,0">
          <category>01</category>
          <group>05</group>
          <symbnum>001</symbnum>
          <variation>01</variation>
          <fill>02</fill>
          <rotation>01</rotation>
        </symb>
      </signbox>
      <txtbox x="19" y="111" char-color="0,0,0">
        <font name="Courier New" size="10" family="74" style="90" weight="92" underlined="0">
          </font>
          <content>LÍNGUA-SINAIS</content>
        </txtbox>
      </cell>
    <cell x="300" y="0" height="150" width="150" bgcolor="255,255,255" cornersColor="0,0,0">
      <signbox>
        <symb x="57" y="24" x-flop="0" y-flop="0" color="0,0,0">
          <category>01</category>
          <group>05</group>
          <symbnum>007</symbnum>
          <variation>01</variation>
          <fill>02</fill>
          <rotation>01</rotation>
        </symb>
        <symb x="58" y="60" x-flop="1" y-flop="0" color="0,0,0">
          <category>02</category>
          <group>03</group>
          <symbnum>012</symbnum>
          <variation>02</variation>
          <fill>01</fill>
          <rotation>01</rotation>
        </symb>
      </signbox>
      <txtbox x="52" y="111" char-color="0,0,0">
        <font name="Courier New" size="10" family="74" style="90" weight="92" underlined="0">
          </font>
          <content>BRASIL</content>
        </txtbox>
      </cell>
    </line>
  </page>

```

Figura 2: Representação SWML do texto LÍNGUA-SINAIS-BRASIL

O programa SignEd [Campos 2003] constitui-se em um editor de textos em SignWriting, capaz de armazenar sinais em um dicionário, e que tem um módulo adicional para geração manual de animações de sinais.

O SignHTML [Campos 2003] é um editor de páginas HTML que utiliza os símbolos do SignWriting em formato GIF para exibição dos textos nas páginas. Tanto o SignEd quanto o SignHTML usam um formato próprio para o armazenamento dos seus dados.

4.3. Dicionários

À parte os dicionários embutidos nos editores de texto, como referido acima, há diversos esforços para a construção de sistemas cuja função específica é o armazenamento de conjuntos de sinais de uma ou de diversas línguas. Esses sistemas são chamados de “dicionários online” pela comunidade do SignWriting.

O principal dicionário on-line disponível hoje em dia é o SignBank [Sutton 2004], desenvolvido por Valerie Sutton, criadora do SignWriting. Além de permitir o armazenamento de sinais, ele possibilita a ordenação dos sinais conforme a ordem lexicográfica própria dos símbolos do sistema SignWriting.

Um uso importante para os dicionários é o apoio ao trabalho em lingüística de sinais, na catalogação dos sinais e de suas características. Nesse caso, uma facilidade importante que o sistema deve ter é a de permitir ao usuário procurar sinais a partir de suas características lingüísticas (fonológicas, etc.). O sistema que está em desenvolvimento na Universidade da Antuérpia [Aerts & Braem 2004] procura atender a esse tipo de finalidade. Esse dicionário também usa SWML extensivamente para armazenamento, importação e exportação de sinais.

4.4. Sintetizadores de Animações de Sinais

Em analogia ao trabalho em síntese de voz, que procura a geração de sons a partir de textos escritos em línguas orais, a síntese de sinais procura a geração de animações de sinais a partir de textos escritos em línguas de sinais.

Ao que sabemos, não há programas capazes de realizar sínteses completas de animações a partir de textos em SignWriting. O principal desenvolvimento nessa área está vinculado à animação de modelos humanos 3D para realidade virtual [Grammalidis 2004].

Um estudo inicial sobre o processo de interpretação da escrita em SignWriting visando a geração automatizada de animações está em curso [Denardi 2003], mas também ainda não dispõe de resultados efetivos.

O Projeto ViSiCast [ViSiCast 2004], por outro lado, avança na direção dessa proposta, mas utiliza o sistema HamNoSys [Hanke 2000] como sistema de escrita, o que restringe o uso do programa a pessoas com conhecimentos especializados em lingüística de sinais, ou a funcionar como módulo de saída para um sistema de tradução automática (que é o propósito final do Projeto ViSiCast).

O programa SignSim [Campos 2003] permite a associação de simulações animadas a sinais e palavras de línguas orais armazenados em dicionários. A exibição de um sinal, por exemplo em um programa de chat, pode ser acompanhada da exibição de sua animação no SignSim. Mas tanto a geração da animação quanto sua associação aos sinais é feita manualmente.

4.5. Reconhecedores Automáticos de Símbolos do Sistema SignWriting

O reconhecimento automático de símbolos do sistema SignWriting é uma das funções mais importantes que precisam ser incorporadas aos sistemas de processamento de línguas de sinais escritas, para agilizar o processo de entrada de textos.

Ainda não se alcançou resultado positivo nessa questão. Em [Rocha 2003], apresenta-se uma proposta de conjunto de símbolos manuscritos para o sistema SignWriting, usando como modelo o conjunto de símbolos do alfabeto “Graffiti” (símbolos alfabéticos usados para escrita manual em computadores do tipo “Palm”).

4.6. Tradutores Automáticos de Sinais baseados em Glosas

Chamam-se “glosas” as traduções simplificadas que se pode fazer entre textos de línguas de sinais e textos de línguas orais, usando um esquema aproximado de tradução em que cada sinal é traduzido por uma palavra ou expressão da língua oral.

Nas Figuras 1 e 3 deste artigo, utilizamos glosas para representar em Português o significado das expressões de línguas de sinais mostradas nas figuras.

Tradutores automáticos baseados em glosas são soluções simples, ainda que precárias, para o problema da tradução automática de línguas de sinais, que aparecem naturalmente quando se cogita uma solução para essa questão.

O SignTalk [Campos 2003] é um desses programas, valendo-se da tradução baseada em glosas para permitir a comunicação entre usuários que se expressam em línguas de sinais e usuários que se expressam em língua oral.

Todo dicionário online ou dicionário incorporado a qualquer um dos editores de texto mencionados acima, ainda que não realizem tradução automática, em essa característica de associar sinais a glosas, para facilitar o entendimento dos sinais por ouvintes.

4.7. Programas para Comunicação via Internet

Juntamente com o desenvolvimento de editores e de dicionários online, o desenvolvimento de ferramentas de comunicação síncrona ou assíncrona via Internet é um dos esforços de desenvolvimento de software para línguas de sinais escritas que tem sido feitos.

O SignTalk [Campos 2003] é um programa de chat que oferece a alternativa de os usuários se comunicarem tanto em uma língua oral quanto em língua de sinais, via escrita em SignWriting. Permite o acoplamento do SignSim (ver abaixo), possibilitando ver animações de sinais.

Um sistema de comunicação assíncrona que está atingindo bom desenvolvimento atualmente é o Sign-WebMessage [Souza 2003]. Em sua versão inicial, dava ao usuário apenas a possibilidade de se comunicar com um conjunto predefinido de símbolos armazenado em um dicionário. A nova versão em desenvolvimento acopla um editor de sinais e permite ao usuário escrever sinais conforme suas necessidades.

Em sintonia com o que acontece com os editores de textos, como o SW-Edit, que visam facilitar aos usuários a geração de páginas HTML, os desenvolvedores dos programas de comunicação via Internet estão em permanente expectativa pela conversão do conjunto de símbolos do sistema SignWriting, de sua atual formatação como mapas de bits (formatos BMP, GIF e PNG, por exemplo) para uma formação vetorial que lhes dê flexibilidade e escalabilidade na apresentação (formato SVG, por exemplo).

A inexistência continuada de conversores de bitmaps para gráficos vetoriais que tenham resolução minimamente adequada às necessidades de representação dos símbolos do SignWriting tem impedido, ainda, a realização dessa conversão.

4.8. Programas de Processamento de Textos Escritos em Línguas de Sinais

Temos conhecimento apenas de dois programas que podem ser caracterizados como programas de processamento de textos escritos em línguas de sinais. Um deles é o procedimento de ordenação de sinais que está embutido no sistema SignBank, o principal dicionário online de sinais escritos em SignWriting. Ele ordena os sinais conforme a ordem lexicográfica dos símbolos do SignWriting. Não é um procedimento que está propriamente aplicado a textos (um dicionário é um banco de sinais, não um texto), mas claramente seu método de ordenação pode ser tomado como base para um procedimento de ordenação dos sinais existentes em um texto.

O outro programa de que temos conhecimento [Costa, Dimuro & Freitas 2004] é o programa demonstrativo do método de procura de sinais em textos de línguas de sinais codificados em SWML, que desenvolvemos. O método tem como elemento central uma técnica de “casamento de padrões de sinais” especialmente definida para levar em conta as pequenas variações na organização espacial dos símbolos, que os usuários podem produzir quando escrevem um sinal. Dando conta dessa variação através de uma particular noção de similaridade entre sinais escritos, o procedimento de casamento de sinais é efetivo e capaz de instrumentalizar qualquer procedimento de procura de sinais em textos de línguas de sinais.

Com a utilização dessa técnica de casamento de sinais codificados em arquivos SWML possibilita que todo tipo de processamento de texto (marcação de partes de discurso, correção ortográfica automática, correção sintática automática, resumo automático, etc.) possa ser feito sobre textos de línguas de sinais.

É claro, a questão da existência de corpora com tamanhos suficientes para que as ferramentas para essas aplicações possam ser desenvolvidas é uma questão crítica, pois a escrita de línguas de sinais é pouco difundida e pouco material escrito foi produzido pelos surdos até agora.

Para superar essa dificuldade, uma das possibilidades com que se deve contar é a da utilização das técnicas desenvolvidas para trabalhar com as chamadas “línguas minoritárias” [Streiter 2003], entre as quais as línguas de sinais podem ser enquadradas, com toda certeza.

No site <http://gmc.ucpel.tche.br/signwriting> pode-se ter acesso a um corpo pequeno de textos escritos em LIBRAS, produzidos por surdos brasileiros.

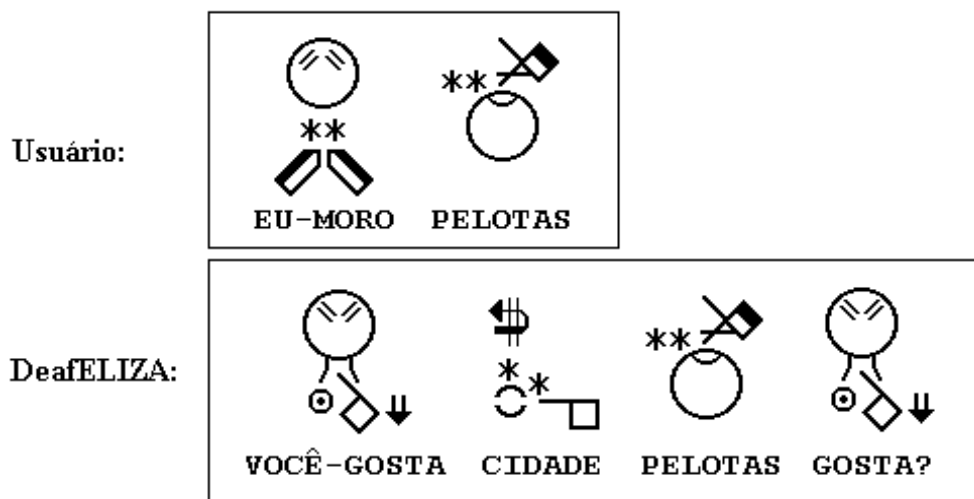


Figura 3: Uma interação em LIBRAS com o programa DeafELIZA

5. Um Programa Elementar de PLS

Esta seção apresenta um exemplo elementar de programa que pode ser propriamente denominado de programa de Processamento de Línguas de Sinais (PLS), no sentido equivalente ao de programa de Processamento de Línguas Naturais (PLN) aplicado a textos em línguas de sinais. Trata-se um programa que desenvolvemos, consistindo de uma simples transposição, para as línguas de sinais, do método de PLN utilizado no histórico programa ELIZA, de Joseph Weizenbaum [Weizenbaum 1966].

Duas foram as razões para a adoção desse modelo de PLN: pela sua simplicidade e pelo seu lado histórico, marcante na área. Não se desconhecem as limitações do modelo, mas também sabe-se da importância histórica que ele teve. Igualmente, tem-se consciência da facilidade de entendimento que ele permite. Por esses dois motivos ele foi escolhido para ser o primeiro modelo de PLN transposto para o processamento de línguas de sinais.

O modelo de PLN de Weizenbaum tem por base dois tipos de procedimentos: procedimentos de análise léxica das frases submetidas ao programa, e procedimentos de casamentos de padrões entre as sentenças analisadas e padrões fixos, previamente embutidos no program.

Em nosso programa, que chamamos de DeafELIZA, adotamos os procedimentos de análise léxica e casamento de padrões de sentenças formulados por Terry Winograd em seu livro-texto [Winograd 1983]. Os procedimentos foram adaptados para se aplicarem aos dados constantes em arquivos SWML, representando textos escritos em SignWriting.

A implementação foi realizada na linguagem de programação Python, contendo apenas o núcleo do programa, isto é, sem que se tenha implementado até agora nenhuma interface gráfica para ele. Assim, as sentenças são submetidas ao DeafELIZA no formato SWML, através de arquivos gerados pelo SW-Edit, e as respostas do programa são mostradas como trechos de textos SWML, em uma janela na tela do computador.

Os padrões de sentenças de línguas de sinais disponíveis até o momento são por demais elementares, sendo apenas suficientes para demonstração das possibilidades técnicas. O DeafELIZA ainda não passou por um desenvolvimento suficiente a ponto de ter alguma possibilidade de uso interativo efetivo com usuários leigos em computação, como acontece com o ELIZA.

A Figura 3 exibe uma interação realizada pelo DeafELIZA, onde tanto a frase submetida ao programa, como a resposta produzida por ele, estão mostradas na forma visual que tem em SignWriting.

6. Conclusão

Os surdos de uma sociedade constituem uma comunidade lingüística e cultural com características próprias que os separam da maioria ouvinte da sociedade [Kyle & Woll 1985, Cuxac 1990, Lane & Philip 1984]. A surdez, deficiência orgânica de caráter funcional, termina produzindo como conseqüência final uma diferença lingüística, cognitiva e cultural entre os surdos e os ouvintes. A comunidade surda de uma sociedade termina desenvolvendo não só sua própria língua, como também sua própria história, seus próprios hábitos, seus valores, sua cultura.

Relacionar linguisticamente uma língua oral (como o Português) com uma língua de sinais (como a LIBRAS) é exatamente como relacionar duas línguas estrangeiras, ainda que elas convivam geograficamente no mesmo país (como o Brasil). Mais ainda, a diferença entre a modalidade oral-auditiva da língua oral e a modalidade gestual-visual da língua de sinais faz com que essa relação seja ainda mais remota do que a relação entre duas línguas orais estrangeiras.

A aplicabilidade das técnicas de PLN às línguas de sinais, como está foi defendida neste artigo, não significa a aplicabilidade imediata às línguas de sinais dos conceitos lingüísticos usualmente utilizados na análise das línguas orais. Significa apenas duas coisas: por um lado, que as características universais das línguas orais se estendem às línguas de sinais, na medida em que essas são línguas naturais exatamente como o são as línguas orais.

Por outro lado, significa que, do ponto de vista operacional, a existência de um sistema de escrita de línguas de sinais prático e destinado ao uso diário, como o sistema SignWriting, permite a transposição direta ao domínio das línguas de sinais das técnicas computacionais normalmente utilizadas para análise de línguas orais. Essa transposição exige, contudo, uma adaptação radical dos módulos de nível lexical-fonológico, para sua adequação à modalidade gestual-visual das línguas de sinais.

No concerne aos níveis lingüísticos mais elevados (estruturas sintagmáticas, estruturas de discurso, etc.) do trabalho de PLN, as especificidades lingüísticas das línguas de sinais não se fazem sentir de modo mais marcante que as especificidades lingüísticas de qualquer língua oral: as variações são de mesma natureza que as encontradas entre as línguas orais e as técnicas de PLN e de Lingüística Computacional podem se aplicar às línguas de sinais sem necessidade de modificações significativas.

Nesse sentido, a tarefa atual da pesquisa em PLS é, precisamente, a de realizar as adaptações necessárias às técnicas de PLN de nível lexical e fonológico, para que elas possam dar o suporte adequado às técnicas de nível mais alto, que podem ser diretamente aplicadas à análise de línguas de sinais. Como este artigo procurou demonstrar, aquela adaptação de técnicas básicas apresenta em si mesmo um interesse técnico-científico elevado.

O artigo apresentou também um convite à comunidade brasileira de PLN para que se envolva com o trabalho em PLS, não só atacando os problemas de transposição de técnicas básicas para o domínio das línguas de sinais, mas também envolvendo-se com os problemas de níveis mais alto. A exigência de conhecimentos sobre as características particulares da LIBRAS, para permitir que esse trabalho possa se realizar, pode estimular a aproximação maior da comunidade nacional de PLN e Linguística Computacional com os poucos lingüistas de línguas de sinais atuantes no país, reforçando com isso a produção de conhecimentos sobre a LIBRAS.

Adicionalmente, os resultados da área de PLS podem ter grande importância para a área de Educação, através do desenvolvimento de sistemas computacionais de apoio ao ensino com interfaces de línguas de sinais, diretamente aplicáveis à Educação Especial de Surdos. Para uma comunidade com elementos geográfica e socialmente dispersos, como a dos surdos, a disponibilidade de tecnologias de informação e comunicação adaptadas à sua especificidade lingüística é, com certeza, um recurso de valor incalculável. A comunidade de PLN pode dar uma contribuição decisiva nesse sentido.

Agradecimentos

A continuidade do Projeto SignNet, desde 1996, só foi possível graças ao apoio financeiro de diversas instituições, agências financiadoras e programas de financiamento à pesquisa. Os autores agradecem à UCPel, CNPq, FAPERGS, UFRGS, PUCRS e ULBRA, Escola Especial Concórdia, Escola Especial Frei Pacífico e Associação dos Surdos de Pelotas. Igualmente, às lingüistas de línguas de sinais Ronice Müller de Quadros e Lodenir Karnopp, pelo suporte lingüístico dado no início do Projeto. A todos os estudantes surdos que participaram do Projeto em diversas épocas. De modo especial, à Marianne Rossi Stumpf, inicialmente estudante secundarista e hoje estudante de doutorado, pelo envolvimento no Projeto desde 1996 e pelo esforço incansável de divulgação da escrita de línguas de sinais, e à Ivana Gomes da Silva e Diogo Souza Madeia, pela atuação na UCPel. À Maria Cristina Pereira, intérprete de LIBRAS, pela participação como pesquisadora do Projeto. À Márcia de Borba Campos, pelo período de parceria. Ao Rafael Pezzin Torchelsen, pela dedicação no desenvolvimento do SW-Edit. À Valerie Sutton e ao Richard Gleaves, pela amizade e apoio permanente. Ao Oliver Streiter, pela receptividade da proposta de PLS na comunidade de processamento de línguas minoritárias.

Referências

- Aerts, S. & Braem, B. (2004) *Site do Dicionário Online das Línguas de Sinais da Bélgica*. Localizado em: <http://pcinf37.ruca.ua.ac.be/aertsbraem/>
- Campos, M. B. *Folder sobre os Programas SignEd, SignSim, SignTalk e SignHTML*. Disponível em: <http://www.inf.pucrs.br/marciabc>
- Costa, A. C. R. *The SWML Site*. Localizado em: <http://swml.ucpel.tche.br>.
- Costa, A. C. R. & Dimuro, G. P. (2001) Supporting Deaf Sign Languages on the Web. *The SignWriting Journal*, v.1, n.0, July 2001. (Available at <http://sw-journal.ucpel.tche.br>. Short version as Poster Paper in the *WWW10 Conference CD-ROM*, Hawaii, 2001).
- Costa, A. C. R. & Dimuro, G. P. (2002) SignWriting-based Sign Language Processing. In: [Wachsmuth & Sowa 1998]
- Costa, A. C. R. & Dimuro, G. P. (2003) SignWriting and SWML: Paving the Way to Sign Language Processing. In: Streiter, O. (Ed.) *Workshop on Minority Languages, Batz-sur-Mer, June 11-14, 2003*. (Satellite workshop of TALN2003 – Traitement Automatique des Langues de Signes).
- Costa, A. C. R.; Dimuro, G. P. & Freitas, J. B. (2004) A Sign Matching Technique to Support Searches in Sign Language Texts. In: Streiter, O. (Ed.) *Workshop on the Representation and Processing of Sign Languages: From SignWriting to .* Lisbon, May 30, 2004. (Workshop associated to LREC2004 - Language Resources 2004).
- Cuxac, C. (1990) *Le Pouvoir des Signes*. In: *Sourds et Citoyens*. Paris: Institut National de Jeunes Sourds de Paris.
- Denardi, R. M. (2003) *Idéias iniciais para a interpretação dos símbolos do Sistema SignWriting visando a geração automática de animações de sinais*. Não publicado. Disponível em: <http://gmc.ucpel.tche.br/swml/animacoes>.
- Glauret, J. (2004) *The ViSiCast Project*. Localizado em: <http://www.visicast.co.uk>.

- Grammalidis, N. (2004) *Gestural Control in 3D Animations using the SignWriting system and SWML*. Localizado em: <http://vsigns.iti.gr:8080/VSigns>.
- Hanke, T. (2004) *HamNoSys – The Hamburg Notation System*. Localizado em: <http://www.sign-lang.uni-hamburg.de/Projects/HamNoSys.html>.
- Kyle, J. & Woll, B. (1985) *Sign Language – The study of deaf people and their language*. Cambridge Univ. Press.
- Lane, H. & Philip, F. (1984) *The Deaf Experience - Classics in Language and Education*. Cambridge: Harvard University Press.
- Maher, J. (1996) *Seeing Language in Sign - The work of William C. Stokoe*. Washington: Gallaudet University Press.
- Quadros, R. M. & Karnopp, L. (2003) *Língua de Sinais Brasileira*. ArtMed, Porto Alegre.
- Rocha, F. Z. (2003) *Proposta de um Padrão Manuscrito para Reconhecimento Automático dos Símbolos do Sistema SignWriting*. Projeto de Conclusão de Curso de Graduação, Escola de Informática/UCPel, 2003. Disponível em: <http://atlas.ucpel.tche.br/fabinha>.
- Souza, V. C. & Crespo, S. (2003) *Sign-WebMessage: Um ambiente para comunicação via Internet baseada na escrita da Língua Brasileira de Sinais*. Mostra de Software do XIV Simpósio Brasileiro de Informática na Educação, Rio de Janeiro, UFRJ/SBC, 12-14/Novembro/2003.
- Stokoe, W. C.; Casteline, D. C. & Croneberg, C. G. (1965) *A Dictionary of American Sign Language on Linguistic Principles*. Gallaudet College Press, Washington.
- Streiter, O. *Example-based NLP for Minority Languages: Tasks, Tools and Resources*. In: Streiter, O. (Ed.) *Workshop on Minority Languages, Batz-sur-Mer, June 11-14, 2003*. (Satellite workshop of TALN2003 – Traitement Automatique des Langues de Signes).
- Sutton, V. (2004) *The SignWriting site*. Localizado em: <http://www.signwriting.org>.
- Sutton, V. (1999) *Lessons in SignWriting - Textbook and Workbook*. La Jolla: Deaf Action Committee for SignWriting. (2nd ed.). Available on-line at: [Sutton 2004].
- Sutton, V. & Gleaves, R. (1995) *SignWriter - The world's first sign language processor*. La Jolla: Deaf Action Committee for SignWriting.
- Sutton, V. (1973) *Sutton Movement Shorthand: a Quick Visual Easy-to-Learn Method of Recording Dance Movement - Book One: The Classical Ballet Key*. Irvine: The Movement Shorthand Society.
- Torchelsen, R. P.; Costa, A. C. R. & Dimuro G. P. (2002) *Editor para textos escritos em SignWriting*. IHC 2002 – 5th. Symposium on Human Factors in Computer Systems, Fortaleza, SBC, 2002. p363–66.
- Torchelsen, R.P. & Costa, A. C. R. (2004) *Programa SWEdit*. Versão atualizada disponível em: <http://gmc.ucpel.tche.br/sw-edit>.
- Valli, C. & Lucas, C. (1995) *Linguistics of American Sign Language - an Introduction*. Washington: Gallaudet University Press. (2nd. ed.).
- Wachsmuth, I. & Sowa, T. (eds.) (2001) *Gesture and Sign Language in Human-Computer Interaction*. Berlin: Springer-Verlag. (LNAI 2298 - Proc. Intl. Gesture Workshop, GW'2001. London, UK, April 18-20, 2001.)
- Weizenbaum, J. ELIZA. *Communications of the ACM*, 9(1966), 36–45.
- Winograd, T. (1983) *Language as a Cognitive Process – Syntax*. Addison-Wesley, Reading.