

See discussions, stats, and author profiles for this publication at: <https://www.researchgate.net/publication/266878408>

SWDB: UM SISTEMA DE DICIONÁRIOS PARA AS LÍNGUAS DE SINAIS USADAS PELOS SURDOS

Article

CITATIONS

0

READS

133

3 authors, including:



Juliano Baldez de Freitas

PROCERGS

5 PUBLICATIONS 14 CITATIONS

[SEE PROFILE](#)



Antônio Carlos Rocha Costa

Pontifícia Universidade Católica do Rio Grande do Sul

313 PUBLICATIONS 1,383 CITATIONS

[SEE PROFILE](#)

Some of the authors of this publication are also working on these related projects:



Formal semantics for social theory [View project](#)



Foundations of Agent Societies [View project](#)

SWDB: UM SISTEMA DE DICIONÁRIOS PARA AS LÍNGUAS DE SINAIS USADAS PELOS SURDOS

Juliano Baldez de Freitas
Universidade Católica de Pelotas
Pelotas – RS – Brasil

Antônio Carlos da Rocha Costa
Universidade Católica de Pelotas
Pelotas – RS – Brasil

RESUMO

O trabalho em andamento pretende desenvolver um sistema de dicionários (léxico) para as línguas de sinais usadas pelos surdos, o SWDB (*SignWriting Data Base*), tem como objetivo proporcionar aos usuários desta língua um auxílio quanto ao significado e as características lingüísticas dos sinais escritos em SignWriting, assim como algumas traduções orais e escritas para estes sinais. A criação deste sistema possibilitará também uma fonte de pesquisa e apoio para futuras aplicações que visem o Processamento da Língua de Sinais.

PALAVRAS-CHAVE

Língua de Sinais, SignWriting, SWML, dicionário, léxico, XML.

1. INTRODUÇÃO

Como toda língua natural, a língua de sinais faz uso de mecanismos computacionais para uma melhor análise de seus aspectos lingüísticos (morfológico, fonológico, sintático e semântico). Através da SignWriting (SW), uma representação gráfica para a escrita da língua de sinais e da SWML (*SignWriting Markup Language*) que é um formato XML para a representação SW, é que se torna possível a análise da língua em seus aspectos lingüísticos. Com a criação de um sistema de dicionários para a língua de sinais é que poderemos armazenar estas informações referentes à língua. O presente artigo irá abordar as ferramentas utilizadas na construção do dicionário. A Seção 2, será abordada a língua de sinais, seu sistema de escrita e a SWML. Na Seção 3 alguns aspectos do processamento da língua natural. A arquitetura do dicionário e protótipo da interface web serão discutidos na Seção 4.

2. LÍNGUA DE SINAIS

Algumas dificuldades são encontradas na educação de pessoas surdas, como por exemplo: ao se comunicarem estas pessoas usam uma língua própria - a língua de sinais -, no entanto são forçadas a se adaptarem a uma língua que não é natural para elas - a língua oral -, tendo enorme dificuldade ao redigir e interpretar as representações escritas da língua oral-auditiva como, por exemplo, o português. Os vários problemas encontrados até então na educação de pessoas surdas motivaram estudos sobre as línguas de sinais utilizadas por essas comunidades surdas. Estes estudos concluíram que a língua de sinais preenche os requisitos lingüísticos que a lingüística coloca para as línguas orais, concluíram, também, que a língua de sinais é uma língua mais acessível aos surdos, constituindo sua língua natural. A língua de sinais é assimilada com rapidez pelos surdos, permitindo uma comunicação eficiente e completa assim como a desenvolvida por ouvintes (Lacerda -2000).

2.1 SignWriting

Pensando na dificuldade de escrita e interpretação da língua oral por parte das pessoas surdas que Valerie Sutton Desta, do *Center For Sutton Moviment Writing*, da Califórnia, USA desenvolveu uma notação gráfica para as línguas de sinais, a SignWriting (Sutton – 1996). A SignWriting consiste em uma representação gráfica da forma gestual da língua de sinais. Ele é um sistema notacional de características gráficas e esquemáticas, constituído de um rico repertório de elementos para a representação dos principais aspectos gestuais das línguas de sinais tais como: configuração de mãos, pontos de articulação, movimentos, expressões faciais, entre outros.

A SignWriting é organizado de forma equivalente à escrita oral, enquanto a escrita oral possui letras e palavras. A SignWriting é constituída de símbolos, aos quais representam um gesto ou movimento na língua de sinais. Um conjunto destes símbolos formam um sinal, o qual forma uma palavra ou um conjunto de palavras (Figura 1).



Figura 1. Representações (sinal e escrita SW) da palavra cachorro.

2.2 SWML

A SWML (*SignWriting Markup Language*) é uma aplicação XML que foi criada por (Costa - 2000) com o intuito de ser um formato a ser utilizado por sistemas que utilizam linguagem de sinais. É um formato para o processamento do SW. Com a SWML é possível o intercâmbio de documentos entre diferentes programas e a análise de textos escritos na língua de sinais, usando a notação gráfica SignWriting independentemente do editor.

3. PROCESSAMENTO DA LINGUAGEM NATURAL

Dentre as diversas áreas da Inteligência Artificial está o Processamento da Linguagem Natural (PLN) que busca compreender os mecanismos da língua (fala e escrita).

Para a criação de uma aplicação que queira obter algum conhecimento da língua, como tradutores, corretores ortográficos, geradores de resumo, etc, são necessárias algumas etapas de análise feita na linguagem, são elas: análise léxica, análise sintática, análise semântica e pragmática (Vieira - 2001).

3.1 Léxico

Uma das etapas descritas acima, a análise léxica, ou também chamada de dicionário é a etapa mais básica e fundamental para o PLN. Esta etapa reúne a estrutura de dados contendo os itens lexicais e as informações correspondentes a estes itens, ou seja, no caso do português, um item lexical seria uma palavra e associada a essa palavra seus mais variados significados além de algumas informações associadas a esta palavra como: categoria gramatical, gênero, número e grau, tempo, regência verbal entre outros (Vieira - 2001). Estas informações são úteis para as etapas posteriores no PLN.

No contexto da língua de sinais o item lexical seria um sinal escrito em SignWriting contendo associados a ele todas as suas características lingüísticas, sinônimos, forma de escrita em SW, código SWML, traduções em línguas orais, etc. A criação de um dicionário deste tipo, aplicado a língua de sinais, é de extrema importância para a melhor compreensão da língua e para a construção de aplicações futuras.

4. DICIONÁRIO SWDB

4.1 XML e banco de dados XML nativo

A XML (*eXtensible Markup Language*) é utilizada para a criação de linguagens próprias de marcações personalizadas, essas marcações descrevem os dados contidos nesta marcação personalizada, isto torna a XML particularmente útil para as áreas cujas informações tenham uma organização complexa e hierárquica (Graves - 2002).

A XML possui características comuns a um banco de dados, por exemplo, ambos armazenam dados, possuem esquemas, linguagens de consulta e interfaces de programação. Os bancos de dados XML nativo possuem algumas vantagens em relação a outros paradigmas quando queremos armazenar dados XML: coleções de documentos, linguagem de consulta, updates e deletes, transações, travas e concorrências, APIs, índices normalização, integridade referencial, escalabilidade. (Graves - 2002).

O uso de um banco de dados para a criação de um sistema de dicionários para a língua de sinais se torna necessário a partir do momento que cada sinal no sistema SignWriting é representado por uma arquivo SWML e possui uma variedade de campos que definem suas características lingüísticas com base em uma língua oral-auditiva, como por exemplo o português, por isso o mais adequado para armazenar esses dados seria um banco de dados XML nativo que trata melhor as questões ligadas a arquivos no formato XML.

4.2 Descrição e arquitetura

O sistema de dicionários SWDB (*SignWriting Data Base*) irá armazenar informações relevantes a um determinado sinal, estas informações são descritas abaixo:

informação visual do sinal (imagem, vídeo, desenho);

representação do sinal em SignWriting (gif, bmp);

código SWML da representação em SignWriting;

espaço para tradução em diversas línguas (português, inglês, etc.);

espaço para anotações lingüísticas, em diversos níveis:

fonológico,

morfológico,

sintático;

semântico.

espaço para exemplos em frases de línguas de sinais usando o sinal;

espaço para sinônimos do sinal em língua de sinais;

espaço para todos esses itens, repetidos, mas escritos em língua de sinais;

espaço para indicação de sinais equivalentes em outras línguas de sinais.

A arquitetura (Figura 2) consiste em um sistema gerenciador (SGBD) com base do sistema 4Suite (4Suite - 2000) que é um pacote do Python (Lutz - 1999) e que contém ferramentas e bibliotecas relacionadas com XML, também funciona como um servidor de documentos XML, com manipulação de dados, tabelas, indexações, o SGBD também terá duas interfaces para o uso “amigável” da ferramenta pelos usuários finais.

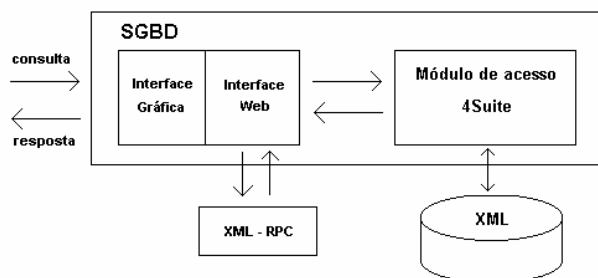


Figura 2. Arquitetura do SWDB.

4.3 Interface web

Este sistema gerenciador do dicionário de sinais terá duas interfaces: uma interface gráfica usando a biblioteca gráfica do Python (wxPython) e outra interface *web* usando XML-RPC (Graves – 2003). Um protótipo da interface *web* está disponível no endereço: <http://swdb.ucpel.tche.br/swdb/>.

A Figura 3 mostra o protótipo da página que faz a busca de sinais através de seu código SWML, ou seja, procura o sinal na base de dados XML através da escrita da língua de sinais. Este é o grande diferencial do SWDB, permite que usuários que não conhecem uma língua oral, procurem informações dos sinais através da sua representação escrita, além de trazer informações lingüísticas dos sinais tanto em línguas orais como de sinais.

Já a Figura 4 mostra o resultado de uma busca feita através do código SWML de um sinal ou através da palavra em uma língua oral. O resultado da busca mostra a representação do sinal em SignWriting, desenho do significado do sinal, um vídeo, uma imagem da representação do sinal, código SWML, sinônimos do sinal, traduções em diversas línguas orais, além de vários aspectos lingüísticos.

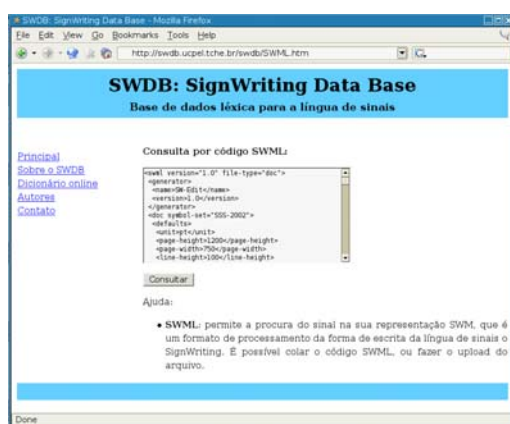


Figura 3. Busca pelo código SWML do sinal.

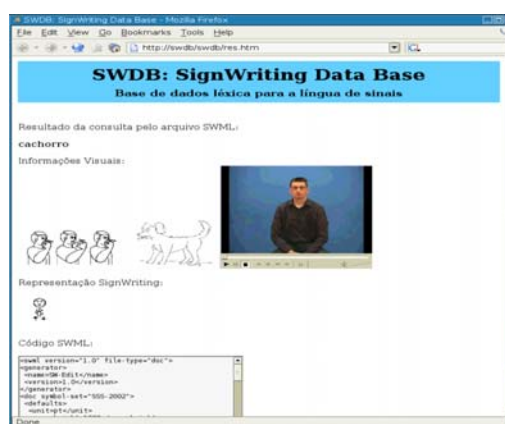


Figura 4. Resultado de uma busca.

5. CONSIDERAÇÕES

O SWDB além de fornecer informações lingüísticas dos sinais usados na língua dos surdos, também tem o objetivo de auxiliar outras ferramentas que visam o processamento da língua de sinais, tais como: analisadores sintáticos, marcadores de partes de discurso, tradutores automáticos, corretores ortográficos, etc.

REFERÊNCIAS

- 4Suite. 4Suite.org. : Fourthought, Inc, 2000. (Available in <http://www.4suite.org/>).
- Costa, A. C. R. SWML - Signwriting Markup Language. Pelotas: ESIN/UCPel, (Available in <http://www.swml.ucpel.tche.br/>).
- Graves, M., 2002 *Projeto de Banco de Dados com XML*. Pearson Books, São Paul, Brasil.
- Lacerda, C. B. F., 2000. A prática pedagógica mediada (também) pela língua de sinais: trabalhando com sujeitos surdos. In Cad. CEDES, page 13. Cad. CEDES.
- Lutz, M. and Acher, D., 1999. *Learning Python*. O'Reilly, Sebastopol, USA.
- Sutton, V. SignWriting: Read, write type Sign Languages. California, USA: The DAC, Deaf Action Committee For SignWriting , 1996. (Available in <http://www.signwriting.org/>).
- Vieira, R. and Lima V. L. S., 2002. Lingüística Computacional: princípios e aplicações. *ERI - ESCOLA REGIONAL DE INFORMÁTICA*, Passo Fundo, Brasil, pp. 210-251.