

**UNIVERSIDADE FEDERAL DE SANTA CATARINA
DEPARTAMENTO DE ENGENHARIA E GESTÃO DO
CONHECIMENTO**

Ronnie Fagundes de Brito

**MODELO DE REFERÊNCIA PARA DESENVOLVIMENTO DE
ARTEFATOS DE APOIO AO ACESSO DOS SURDOS AO
AUDIOVISUAL**

Tese de doutorado submetida ao Programa de Engenharia e Gestão do Conhecimento da Universidade Federal de Santa Catarina para a obtenção do Grau de Doutor em Engenharia e Gestão do Conhecimento
Orientador: Prof.^a Alice Theresinha Cybis Pereira, PhD.

Florianópolis

2012

Ficha de identificação da obra elaborada pelo autor,
através do Programa de Geração Automática da Biblioteca Universitária
da UFSC.

Brito, Ronnie Fagundes de

Modelo de Referência para Desenvolvimento de Artefatos
de Apoio ao Acesso de Surdos ao Audiovisual [tese] /

Ronnie Fagundes de Brito ; orientadora, Alice Theresinha
Cybis Pereira - Florianópolis, SC, 2012.

337 p. ; 21cm

Tese (doutorado) - Universidade Federal de Santa
Catarina, Centro de Comunicação e Expressão. Programa de Pós-
Graduação em Engenharia e Gestão do Conhecimento.

Inclui referências

1. Engenharia e Gestão do Conhecimento. 2.
Acessibilidade. 3. Surdos. 4. Audiovisual. I. Pereira,
Alice Theresinha Cybis . II. Universidade Federal de Santa
Catarina. Programa de Pós-Graduação em Engenharia e Gestão
do Conhecimento. III. Título.

Ronnie Fagundes de Brito

**MODELO DE REFERÊNCIA PARA DESENVOLVIMENTO DE
ARTEFATOS DE APOIO AO ACESSO DOS SURDOS AO
AUDIOVISUAL**

Esta tese foi julgada adequada para obtenção do título de “Doutor em Engenharia e Gestão do Conhecimento”, e aprovada em sua forma final pelo Programa de Engenharia e Gestão do Conhecimento.

Florianópolis, 24 de julho de 2012.

Prof. Paulo Maurício Selig, Dr.Eng.
Coordenador do Curso

Banca Examinadora:

Prof.^a Alice Theresinha Cybis Pereira,
PhD.
Universidade Federal de Santa Catarina
Orientadora

Prof. Fernando Ostuni Gauthier,
Dr.Eng.
Universidade Federal de Santa Catarina
Membro

Prof. José Leomar Todesco, Dr.Eng.
Universidade Federal de Santa Catarina
Membro

Prof. José Mario De Martino, Dr.Eng.
Universidade Estadual de Campinas
Examinador Externo ao Programa

Prof.^a Ronice Müller de Quadros, Dr.^a
Universidade Federal de Santa Catarina
Examinador Externo

Prof. Vilmar Silva, Dr.
Instituto Federal Santa Catarina
Examinador Externo

Ronnie Fagundes de Brito

**MODELO DE REFERÊNCIA PARA DESENVOLVIMENTO DE
ARTEFATOS DE APOIO AO ACESSO DOS SURDOS AO
AUDIOVISUAL**

Prof. Tarcísio Vanzin.
Universidade Federal de Santa Catarina
Moderador

*

AGRADECIMENTOS

A todos que de alguma forma contribuíram nesse trabalho.

Especificamente à equipe do Hiperlab, Alice Cybis, Marilinha, Berenice, Bruno Velloso, Mônica Renneberg, Robson Lorberski, Nícolas Timmermann, Natalia Cardoso, Michel Kramer, João Garcia, Felipe Gondim, Gabriel Fronza e outros que incentivaram e apoiaram este desenvolvimento.

À Ronice Quadros, Marianne Stumpf, Janine Soares, Rodrigo Marques, Uéslei Paterno, Silvana Nicoloso, Silvana Aguiar, Shirley Vilhalva, Kafka, Vinicius, Rundesth Saboya, Marcos Brito, Vanessa Amadeo, Rafael Marques e demais colegas do curso de bacharelado e licenciatura em Letras-Libras.

Aos professores e colegas do EGC e EGR, Fialho, Vanzin, Neri dos Santos, Roberto Pacheco, Vânia Ulbricht, Marília Matos, Fernando Spanhol, Marta Braga, Rosane Obregon, Emilio Brkanitch, Douglas Kaminsky, Kamil Giglio, Greyci Spanhol e outros que promoveram enriquecedoras discussões sobre temas relacionados às mídias, engenharia e gestão do conhecimento.

Aos alunos da UFSC e da FEAN, que tanto desafios propuseram durante o tempo dedicado à pesquisa.

Aos amigos Rodolfo Ramos, Joel Prosdócimo, Robson Costa, Ângela Araújo, Fernando Sing, Allan Schiestl, Jonatas Carvalho, Jefferson Silva, Leandro Hommerding, Michel Schwingel, Fábio Mielke, e inúmeros outros que compartilharam momentos de alegria e convivência confortante durante o período dedicado a esta pesquisa.

Aos membros da banca, os professores Vilmar Silva, incentivador dos estudos relacionados aos Surdos, José de Martino, pesquisador de recursos associados à acessibilidade de surdos, Fernando Gauthier e José Todesco, grandes professores desde os tempos da graduação.

E finalmente, a meus pais, Marli e Werner, e a meus irmãos, Renne e Renata, por terem a disposição em me apoiar e sem os quais nada disso seria possível.

"Deficiente" é aquele que não consegue modificar sua vida, aceitando as imposições de outras pessoas ou da sociedade em que vive, sem ter consciência de que é dono do seu destino.

"Louco" é quem não procura ser feliz com o que possui.

"Cego" é aquele que não vê seu próximo morrer de frio, de fome, de miséria, e só tem olhos para seus míseros problemas e pequenas dores.

"Surdo" é aquele que não tem tempo de ouvir um desabafo de um amigo, ou o apelo de um irmão. Pois está sempre apressado para o trabalho e quer garantir seus tostões no fim do mês.

"Mudo" é aquele que não consegue falar o que sente e se esconde por trás da máscara da hipocrisia. [...]

Autor desconhecido

RESUMO

As tecnologias da informação e comunicação possibilitam a participação do sujeito na sociedade do conhecimento, entretanto o tema da acessibilidade dos surdos aos conteúdos audiovisuais em meios digitais ainda demanda estudos para viabilizar sua efetiva e ampla adoção.

Objetiva-se identificar e analisar as alternativas para o desenvolvimento de um modelo de referência que oriente o reuso de processos, métodos e técnicas para a produção de artefatos que promovam a acessibilidade de surdos aos conteúdos audiovisuais em plataformas digitais.

A partir de uma revisão sistemática da literatura são apontadas recomendações para apresentação de conteúdo audiovisual acessível ao público surdo, os requisitos que devem ser atendidos para promover as estratégias de acesso utilizadas por diferentes perfis de surdos, e enumeradas alternativas que podem apoiar estas demandas, como o uso de legendas textuais e com janela de língua de sinais.

O modelo de referência contempla a produção de conteúdos a partir da tradução do material audiovisual, sendo identificadas e elaboradas recomendações para a geração de legendas em vídeo de língua de sinais ou na forma escrita. Busca-se integrar a produção destes tipos de artefatos, por meio de processos manuais ou automáticos sendo identificadas as mídias que apoiam ou são resultantes dos processos de produção de artefatos de apoio a acessibilidade.

O modelo de referência é validado diante a consulta a especialistas e aplicado em uma implementação de referência de um sistema para acessibilidade com cenários de entrega na televisão digital interativa e na web.

Como resultados são apresentadas as recomendações e alternativas em relação aos processos e mídias necessárias para a acessibilidade dos surdos ao audiovisual digital.

Palavras-chave: Surdos; Acessibilidade; Audiovisual.

ABSTRACT

Information and communication technologies allow people participation in knowledge society, however accessibility issues related to deaf and audiovisual content in digital media still requires studies to facilitate its effective and widespread adoption.

The objective of this research is to develop a reference model that identifies and analyze alternatives for the reuse of processes, methods and techniques related to the production of artifacts that promote accessibility for deaf in audiovisual content on digital platforms.

From a systematic review of the literature, recommendations for presentation of audiovisual content to deaf people and the requirements that must be met to promote access strategies used by different deaf subjects are pointed, and alternatives that can support these demands, as the use of textual captions and sign language window, are listed and characterized.

The reference model includes the production of content from the translation of audiovisual material, being identified and developed recommendations for the generation of video subtitles or sign language in the written form. It attempts to integrate the production of these types of artifacts, through manual or automated processes are identified that support the media or are the result of the production processes of artifacts to support accessibility.

The reference model is validated on the expert consultation, and applied to a reference implementation of a system for accessibility to delivery scenarios in interactive digital television and the web.

As results, the recommendations in relation to media alternatives and processes necessary for the accessibility of the digital audiovisual by deaf people are presented.

Keywords: Deaf; Accessibility; Audiovisual.

LISTA DE FIGURAS

Figura 1: Exemplo de janela com tamanho insuficiente	30
Figura 2: Espaço reservado à apresentação da Libras de acordo com recomendação de acessibilidade.	31
Figura 3: Estrutura base de um processo Delphi	39
Figura 4: Manifestação em prol da educação bilíngue.....	53
Figura 5: Material de ensino de línguas de sinais na Escola Estadual de Educação Especial Dr. Reinaldo Coser.....	54
Figura 6: Exemplo de material didático para alfabetização de surdos	55
Figura 7: Três notações para o registro escrito das línguas de sinais.	60
Figura 8: Texto no sistema de escrita Elis.	61
Figura 9: Tradução livre para o português do texto em LS.....	61
Figura 10: exemplos de configurações de mão e sua notação correspondente em HamNoSys.....	62
Figura 11: exemplo de correspondência entre sinais e escrita de sinais.....	62
Figura 12: Correspondência entre configuração de mão e seu símbolo em escrita de sinais.	63
Figura 13: correspondência de símbolos com expressões faciais referentes às bochechas.	63
Figura 14: Exemplo de ordem de leitura/soletragem de um sinal. Os símbolos à direita estão ordenados pela ordem de leitura.	66
Figura 15: área de web site de emissora dedicada à oferta de conteúdo acessível.	68
Figura 16: TILS em noticiário francês.	69
Figura 17: TILS acompanhado de legenda textual.	69
Figura 18: iniciativa de TV interativa na web “surd@ TV Brasil”	70
Figura 19: interface web para apresentação de vídeo com opção transcrição automática de áudio para geração de legendas em texto.....	71
Figura 20: exemplo de aplicação de modulo de apresentação de LS	72
Figura 21: telejornal dedicado aos surdos usuários de Libras	73
Figura 22: Sistema de TV digital.....	74
Figura 23: modelo de metadados para serviços de TV interativa	76
Figura 24: pirâmide de <i>Vauquouis</i> adaptada por Januário, Leite, Koga (2010) ...	80
Figura 25: exemplo de problema com legendagem em <i>closed-captions</i> analógico	91
Figura 26: exemplos de apresentação da língua de sinais em conteúdo audiovisual na TVDI.....	100
Figura 27: Exemplo DVD da Monica "Uma Aventura no Tempo"	100
Figura 28: Exemplo vídeo do YouTube.....	101
Figura 29: Exemplo de legenda com Libras em conteúdo ao vivo pela web e TV	101
Figura 30: totalidade do intérprete de Libras com o programa de TV	109

Figura 31: Diagrama de atividades para o processo para produção de legendas em vídeo.	110
Figura 32: Quadro do vídeo original e quadro do vídeo com aplicação de algoritmo de extração de contornos.	113
Figura 33: comparação entre codificações de vídeo a 55kbps.	114
Figura 34: combinação de resolução espacial e taxa de quadros	116
Figura 35: Aplicativo ELAN utilizando glosas para transcrição do conteúdo em língua de sinais.	119
Figura 36: Visualização de mensagem em escrita de sinais em dispositivo móvel	120
Figura 37: Opinião de surdos sobre uso de software de comunicação por escrita de sinais.	121
Figura 38: Web site em escrita de sinais.	122
Figura 39: Proposta de aplicação de legenda com escrita de sinais.	123
Figura 40: Editor de texto em escrita de sinais	124
Figura 41: Editor de texto de sinais SignText.	125
Figura 42: Editor de texto de sinais SW-Edit.	126
Figura 43: Interface de cadastro de sinal em dicionário.	127
Figura 44: Interface de consulta em dicionário a partir de escrita de sinais. ...	128
Figura 45: Interface de busca em dicionário a partir de escrita oral.	128
Figura 46: Descrição do termo "Ambiente Virtual" no glossário Letras-Libras.	129
Figura 47: dinâmica de desenvolvimento e uso de glossário em LS	130
Figura 48: interface de busca de sinais a partir de cheremas representados por avatar	131
Figura 49: Interface com parâmetros de busca em LS.	132
Figura 50: Resultado da busca por sinais.	132
Figura 51: Lista de dicionários em língua de sinais SignBank.	133
Figura 52: Pesquisa por palavra em dicionário de LS.	134
Figura 53: arquitetura de sistema de reconhecimento de voz para geração de vídeo em língua de sinais.	141
Figura 54: Estrutura Funcional do FALibras	142
Figura 55: Arquitetura do módulo de tradução do FALibras MT	143
Figura 56: exemplo de agente embutido em sistema de dialogo falado.	144
Figura 57: Visão geral de sistemas de sinais que utilizam avatares	146
Figura 58: Exemplo de um avatar e um semantróide.	147
Figura 59: pontos de articulação da especificação H-anim.	150
Figura 60: agente 3D interpreta a frase "Eu vou para a escola de manhã" em LS vietnamita.	152
Figura 61: exemplo de SML(X)	152
Figura 62: Interface para animação direta dos sinais em um avatar	153
Figura 63: arquitetura de um serviço de geração de vídeos com avatares sinalizantes	154
Figura 64: Arquitetura de Web Service para síntese de sinais em CSL (Chinese Sign Language).	155

Figura 65: exemplo de HamNoSys e SigML	157
Figura 66: exemplo de SignWriting e SWML	157
Figura 67: sequência para animação de sinais com H-Anim a partir de SignWriting	159
Figura 68: exemplo de animação de expressão facial a partir de escrita de sinais.	161
Figura 69: Avatar VGuido	162
Figura 70: avatar do SignSynth	165
Figura 71: Agente de representação de gestos.	166
Figura 72: Software Rybená apresentando legenda em LS	167
Figura 73: visão geral do sistema de animação de gestos de sinais AGA-Sign	168
Figura 74: sequência de animação de sinais a partir de SW	168
Figura 75: Arquitetura do sistema de tradução PoliLibras	171
Figura 76: Virtual Jonah apresentando LS	173
Figura 77: avatares Robby e Torrents renderizados e seus sistemas de esqueletos.	175
Figura 78: configurações de mão realizadas por um sinalizante humano e por avatar	175
Figura 79: Modelo de casos de uso para a produção de artefatos de apoio.	185
Figura 80: Modelo de casos de uso para o acesso do surdo ao audiovisual.	186
Figura 81: Hierarquia de classes de artefatos para acessibilidade ao audiovisual.	188
Figura 82: Tipos de técnicas para produção de legendas com LS	190
Figura 83: Hierarquia de classes para animação de avatares.	191
Figura 84: Diagrama de classes do glossário.	192
Figura 85: Sequência de atividades para produção de legendas em texto.	194
Figura 86: Sequência de atividades para produção de legenda com vídeo.	196
Figura 87: Atividades para produção de legendas em escrita de sinais.	197
Figura 88: Atividades para produção de legenda de LS com avatares.	199
Figura 89: Legenda para modelo de alternativas de processos para produção de legendas.	201
Figura 90: Modelo de alternativas de processos e artefatos para produção de legendas	202
Figura 91: Modalidades de comunicação e possíveis traduções para desenvolvimento de audiovisual acessível aos surdos.	203
Figura 92: Modelo de casos de uso para produção de artefatos de apoio atualizado.	249
Figura 93: Módulos da arquitetura utilizada na implementação.	251
Figura 94: Exemplo de código para legenda em escrita de sinais.	254
Figura 95: Interface do software para composição de vídeos Moviemasher.	255
Figura 96: Interface do Moviemasher adaptada à tarefa de tradução/ composição de audiovisual acessível para operação de sistema desenvolvido a partir do modelo de referência proposto.	256
Figura 97: Exemplo de dados para legendas em texto oral	257

Figura 98: Exemplo de dados para legendas em LS com vídeo	258
Figura 99: Exemplo de dados para legendas em escrita de sinais.....	258
Figura 100: Exemplo de dados para mídia de vídeo.....	258
Figura 101: Integração do editor de texto de sinais ao editor de legendas.....	260
Figura 102: modelo para apresentação de legenda em LS na forma escrita	262
Figura 103: Código de definição das regiões de exibição das legendas na TV digital.....	264
Figura 104: Código para definição de medias para acessibilidade de surdos na TV digital.	264
Figura 105: Elementos de modelo de referência utilizados no cenário de uso.	268
Figura 106: Elementos de modelo de referência utilizados no cenário de uso.	271
Figura 107: Editor de texto SW-Edit.	272
Figura 108: interface do aplicativo de produção de legendas <i>Subtitle Workshop</i>	274
Figura 109: Exemplo de fruição de conteúdo com legenda em escrita de sinais na plataforma de TV digital.....	284
Figura 110: Elementos do modelo utilizados no cenário de uso.....	285
Figura 111: Transcrição bilíngue criada por Adam Frost.	286
Figura 112: Elementos do modelo utilizados no cenário de uso.....	287
Figura 113: Exemplo de diferença na disposição dos símbolos manuais quando usada a datilologia.	295
Figura 114: Transcrição do conteúdo traduzido em LS.	333
Figura 115: Transcrição do conteúdo traduzido em LS.	335
Figura 116: Caracterização dos especialistas consultados	337

LISTA DE QUADROS

Quadro 1: Número de pessoas com dificuldades de audição	35
Quadro 2: categorias de símbolos do <i>SignWriting</i>	66
Quadro 3: A natureza polissemiótica dos produtos audiovisuais.	77
Quadro 4: modalidades de tradução para o acesso do surdo ao audiovisual e artefatos resultantes	85
Quadro 5: Classificação da legendagem em português de acordo com os parâmetros de Teixidor (2008).....	93
Quadro 6: Classificação de tecnologias para produção de conteúdo em LS	104
Quadro 7: comparativo de técnicas para a síntese de conteúdo em línguas de sinais.	136
Quadro 8: Comparação entre diferentes artefatos para comunicação em LS ...	147
Quadro 9 : variação distal proximal na representação de entonação por avatar.	177
Quadro 10: Análise das tecnologias para produção de LS partir da taxonomia de Fajardo, Vigo e Salmeron (2009).....	180
Quadro 11: Mapeamento de iniciativas relacionadas ao modelo	203
Quadro 12: Formatos para apresentação de legendas para surdos	252
Quadro 13: Modelo para interação com aplicativo de exibição de legendas. ...	263
Quadro 14: Exemplo de diferentes níveis de condensação de legendas	274
Quadro 15: Correspondência entre conteúdo da transcrição em língua oral e em LS	275
Quadro 16: quantidade de artigos resultantes da revisão sistemática.....	325
Quadro 17: quantidade de teses e dissertações resultantes da revisão sistemática	326
Quadro 18: Validação estatística das recomendações na primeira rodada.	331
Quadro 19: Validação estatística das recomendações na segunda rodada.	332
Quadro 20: Descrição dos especialistas consultados	337

LISTA DE ABREVIATURAS E SIGLAS

ASL – *American Sign Language*
BAML - Body Animation Markup Language
BAP – Body Animation Parameters
CIF - Common Interchange Format
EML - Emotional Markup Language
EPG – Electronic Program Guide
FAP – Face Animation Parameters
FPS - frames per second
FAML - Facial Animation Markup language
GML - Gesture Markup Language
H-Anim – Humanoid Animation
Kbps – Kilobytes por Segundo
Libras – Língua Brasileira de Sinais
LO – Língua Oral
LS – Língua de Sinais
LSF – *Langue des Signes Français*
DGS – *Deutsch GebärdenSchrift*
NCL –Nested Context Language
P2P – Peer to Peer
PVR - Personal Video Recorder
QCIF - Quarter CIF
SigML – Sign Gesture Markup Language
SML - Speech Markup Language
SQCIF – Sub QCIF
SRT – Subrip
SW- SignWriting
SWML – SignWriting Markup Language
TSWML – Timed SignWriting Markup Language
TTL – Timed Text Language
TVD – Televisão Digital
TVDI – Televisão Digital Interativa
TVML – Television Markup Language
VHML – Virtual Human Markup Language
VRML – Virtual Reality Markup Language

SUMÁRIO

1 Introdução.....	27
1.1 <i>Contextualização e Definição do Problema</i>	27
1.2 <i>Questão de pesquisa</i>	32
1.3 <i>Objetivos</i>	32
1.4 <i>Justificativa</i>	33
1.5 <i>Metodologia da pesquisa</i>	36
1.5.1 <i>A visão de mundo</i>	36
1.5.2 <i>Procedimentos metodológicos</i>	38
1.6 <i>Originalidade</i>	41
1.7 <i>Escopo e Limites da pesquisa</i>	41
1.8 <i>Aderência ao programa</i>	42
1.9 <i>Resultados esperados</i>	43
1.10 <i>Estrutura da tese</i>	43
1.11 <i>Envolvimento pessoal com a pesquisa</i>	44
1.12 <i>Resumo do capítulo</i>	45
2. Acessibilidade ao audiovisual: os Surdos, a Tradução e Ambientes de Acesso.....	46
2.1 <i>O público surdo</i>	47
2.1.1 <i>Uma diversidade de visões</i>	47
2.1.2 <i>A Educação Bilíngue</i>	53
2.1.3 <i>As Línguas de Sinais</i>	56
2.1.4 <i>Escrita das Línguas de Sinais</i>	59
2.2 <i>Ambientes de Acesso</i>	68
2.3 <i>A tradução do audiovisual para os surdos</i>	77
2.3.1 <i>Tradução e Interpretação</i>	78
2.3.2 <i>Modalidades de tradução aplicadas ao audiovisual</i>	82
2.4 <i>Conclusões sobre o acesso de surdos ao audiovisual, Ambientes de Acesso e Tradução</i>	87
3. Artefatos de apoio ao acesso dos surdos ao audiovisual	89
3.1 <i>Legendas em Texto na Língua Oral</i>	90
3.1.1 <i>Produção de Legendas Textuais em Língua Oral</i>	94
3.2 <i>Legendas com Intérprete de Línguas de sinais</i>	99
3.2.1 <i>Produção de Legendas com Intérprete de Libras com Vídeo</i> ..	106
3.2.1.1 <i>Codificação de Vídeo em Línguas de Sinais</i>	112
3.3 <i>Legendas de Texto em Língua de sinais</i>	119
3.3.1 <i>Produção de Legenda em Escrita de Sinais</i>	124
3.4 <i>Glossários de Línguas de Sinais</i>	127
3.5 <i>Sistemas de tradução automática</i>	135

3.5.1 Reconhecimento das Línguas de Sinais a partir de vídeo em LS	137
3.5.2 Geração de LS a partir de Vídeos Pré-gravados	141
3.5.3 Síntese de Línguas de Sinais com Intérpretes Virtuais.....	144
3.5.3.1 Padrões para dados de animação de avatares	149
3.5.3.2 Síntese concatenativa	151
3.5.3.3 Síntese Articulatória.....	157
3.5.3.4 Avatares bidimensionais	166
3.5.3.5 O software de tradução PoliLibras	170
3.5.3.7 Questões relativas à qualidade da síntese de sinais em avatares.....	174
3.5.3.8 Conclusões sobre síntese de sinais a partir de avatares	178
3.6 <i>Caracterização das tecnologias para produção e apresentação de línguas de sinais</i>	179
3.7 <i>Conclusões sobre artefatos de apoio à acessibilidade e ambientes de acesso</i>	182
4. Proposta de modelo de referência	184
4.1 <i>Modelos de Casos de uso para produção e acesso ao audiovisual</i> ..	185
4.2 <i>Modelos de classes</i>	188
4.2.1 Modelo de classes de artefatos para acessibilidade.....	188
4.2.2 Modelo de Técnicas de Produção de Legendas Gestuais	190
4.2.3 Modelo de classes para o glossário	192
4.3 <i>Modelos de atividades dos processos para produção de artefatos</i> ..	193
4.3.1 Modelo de atividades para produção de legendas em texto oral	194
4.3.2 Modelo de Atividades para Produção de Legendas de Sinais com Intérprete Real.....	196
4.3.3 Modelo de Atividades para Produção de Legendas em Escrita de Sinais	197
4.3.4 Modelo de atividades para produção de Legendas em Língua de Sinais com Intérprete Virtual	198
4.3.5 Modelo Integrado das Atividades para Produção de Artefatos de Apoio ao Acesso dos Surdos ao Audiovisual	200
4.4 <i>Descrição do modelo de acordo com a visão de mundo</i>	205
4.5 <i>Considerações sobre o modelo</i>	205
5. Questionário Delphi	206
5.1 <i>Surdos e audiovisual</i>	208
5.2 <i>Papéis</i>	213
5.3 <i>Artefatos</i>	222
Comentários sobre Artefatos de Apoio	226

5.4	<i>Produção de legenda em português escrito</i>	227
	Comentários sobre Produção de Legendas em Português Escrito	230
5.5	<i>Produção de legenda de vídeo em Língua de Sinais</i>	232
5.6	<i>Produção de legendas em LS com avatares</i>	236
	Comentários sobre Produção de Legendas com Avatares	239
5.7	<i>Produção de legenda em escrita de sinais</i>	241
5.8	<i>Validação estatística</i>	247
5.9	<i>Conclusões sobre a consulta aos especialistas</i>	248
6.	Implementação de referência e Cenários de Uso	250
6.1	<i>Implementação de Referência</i>	251
6.1.1	Editor de legendas	255
6.1.1.1	Esquemas de dados do editor de legendas	257
6.1.2	Gravador de Vídeo e Teleprompter	259
6.1.3	Importação de Legendas	259
6.1.4	Editor de Legendas de Sinais	260
6.1.5	O dicionário/glossário SignWriting	261
6.1.6	Aplicativo para exibição de legendas	262
6.2	<i>Cenários de uso</i>	266
6.2.1	Cenário de uso com vídeo educativo na web	266
6.2.1.1	Produção	266
6.2.2	Cenário de Telenovela na TVD	269
6.2.2.1	Produção	269
6.2.3	Cenário de Previsão do Tempo na Web	272
6.2.3.1	Produção	272
6.2.3.2	Fruição	284
6.2.4	Cenário de transcrição de vídeo em LS.....	286
	<i>O próximo tópico apresenta as conclusões a respeito do processo para a implementação de referência e dos cenários de uso</i>	287
6.3	<i>Considerações sobre a implementação de referência e cenários de uso</i>	288
7.	Conclusões	289
7.1	<i>Pesquisas futuras</i>	294
7.2	<i>Considerações Finais</i>	297
	REFERÊNCIAS	298
	ANEXO A – Revisão sistemática da bibliografia	325
	ANEXO B – Validação Estatística das respostas ao questionário DelpHi	331
	ANEXO C – Transcrições do conteúdo em escrita de sinais	333
	ANEXO D – Caracterização dos especialistas consultados	337

1 INTRODUÇÃO

O acesso à informação é um dos fatores essenciais para participação na sociedade do conhecimento, que utiliza intensamente as tecnologias de comunicação digitais. O exemplo mais significativo de tecnologia digital é a rede Internet, e mais recentemente, a televisão. Esta última sempre teve papel importante na sociedade, informando seus telespectadores e lhes proporcionando entretenimento. Com sua digitalização a TV encontra potencial para aprimorar seus serviços e conteúdos, oferecendo entre outros, maior acessibilidade.

Em relação à acessibilidade deve se considerar diferentes requisitos de acordo com os diferentes perfis de usuários, observando-se aspectos como a constituição da percepção do sujeito, o contraste de cores e a legendagem (CAÑON, 2005).

Entre usuários surdos, nem todos possuem nível de alfabetização no português adequado à leitura de legendas, e nem todos se comunicam por meio da Língua Brasileira de Sinais (Libras), de forma que se deve disponibilizar a legenda tanto em português quanto em Libras, sempre oferecendo acesso visual à informação sonora (TORRES; MAZZONI; ALVES, 2002).

Cabe ressaltar que as línguas de sinais estão sub-representadas no mundo virtual, fato que constitui uma barreira de acesso para a comunidade surda (FAJARDO; VIGO; SALMERON, 2009).

1.1 CONTEXTUALIZAÇÃO E DEFINIÇÃO DO PROBLEMA

O audiovisual é uma forma de comunicação de massa que até os dias atuais tem influenciado fortemente a cultura brasileira. A TV como fonte de informação, mesmo que parcial, situa o sujeito na sociedade e o coloca a par dos acontecimentos locais, nacionais e mundiais. A compreensão das mensagens enviadas pela TV determina aspectos da vida cotidiana, como a roupa que se veste ou o assunto das conversas de corredor. Além disso, o audiovisual é meio para o entretenimento e a educação.

No cenário do conteúdo audiovisual, a televisão digital interativa (TVDI) surgiu como forma de proporcionar imagens de maior qualidade e interatividade a seus espectadores (MONTEZ; BECKER, 2005), e permitir a disseminação de conhecimentos na sociedade em rede (BALDESSAR; GIGLIO, 2010). Nesse contexto, a acessibilidade é recurso para a inclusão digital de diferentes públicos. Deve-se

considerar desde a viabilidade do acesso das pessoas aos recursos computacionais que vão apresentá-lo, quanto o acesso aos conteúdos oferecidos. A importância do acesso dos surdos aos conteúdos em ambientes digitais é destacado por Corradi e Vidotti (2007):

A acessibilidade está intrinsecamente relacionada à usabilidade de ambientes digitais por usuários surdos que, por meio de conteúdos [...] com a presença da Libras e com legendas em português para vídeos com áudio, podem encontrar possibilidades inclusivistas e de participação efetiva na Sociedade da Informação e do Conhecimento (CORRADI; VIDOTTI, 2007, p.13).

Neste sentido, a televisão pode “proporcionar uma enriquecedora experiência bilíngue para os Surdos e para os ouvintes” (SOUZA, 2005, p.124). A TV também tem papel importante na formação de uma sociedade mais inclusiva ao oferecer acesso à mídia de massa, a qual tem importância destacada quando informações de interesse geral são disseminadas (KURZ; MIKULASEK, 2004). Por outro lado,

A televisão ao passar imagens sem legenda ou tradução para os surdos está constituindo o lugar da incapacidade de comunicação. Os surdos diante da dificuldade de atribuir sentidos para o que estão vendo na TV, ficam subjugados aos ouvintes que detém a informação e antes disso, que detém a capacidade de ouvir (Reichert, 2006, p. 29).

A escolha de quais artefatos e estratégias¹ os surdos utilizarão para ter acesso ao conteúdo audiovisual depende de fatores como a cultura educacional, o contexto familiar e o momento da surdez. Essa diferença é exemplificada por (KENNAWAY; GLAUERT; ZWITSERLOOD, 2007):

[...] aqueles que perderam a audição tarde na vida podem facilmente acessar a informação textual. No entanto, para os indivíduos surdos pré-linguísticos (aqueles que eram surdos desde antes

¹ Estes conceitos são definidos ao decorrer do trabalho, em específico na sessão 1.5.1 que trata da visão de mundo adotada na pesquisa.

de aprender qualquer língua), o material escrito muitas vezes é muito menos acessível do que se apresentado em língua de sinais (KENNAWAY; GLAUERT; ZWITSERLOOD, 2007, p.16, tradução nossa).

Souza (2005), ao entrevistar um grupo de surdos afirma que “para determinados surdos é muito difícil compreender os assuntos apresentados na televisão por meio de imagens, mesmo acompanhadas de legendas”. Um dos entrevistados surdos afirma que:

Para os Surdos a legenda é importante, o intérprete também é importante. O Surdo sabe ler, sabe um pouco o português, a legenda é rápida, dá confusão. Na Globo, o Jornal Nacional tem legenda, mas falta intérprete, e nas novelas não tem legenda e não tem intérprete, é feita só para ouvintes (SOUZA, 2005, p.122).

Apesar de atualmente as legendas estarem mais presentes na programação televisiva, desde o início das transmissões televisivas existe pouca presença de legendas em línguas de sinais, pois o Tradutor e Intérprete de Língua de Sinais (TILS) utiliza um espaço na tela que não é atrativo para audiências ouvintes (DELGADO, 2006), e sabe-se que o uso de legendas textuais é voltado principalmente à adaptar conteúdos estrangeiros à língua pátria.

Existem poucas análises e propostas de soluções em relação às legendas que são transmitidas para a população brasileira atualmente (SELVATICI, 2010). Na TV analógica, a janela de Libras, quando presente, ocupa pequena parte da tela, onde em muitos casos a janela não é grande o suficiente para permitir a leitura adequada dos sinais (PICCOLO; MELO; BARANAUSKAS, 2007).

A janela, às vezes, é tão pequena que não é possível ver as mãos do intérprete e, muito menos, é possível distinguir os sinais que este está fazendo (REICHERT, 2006, p. 61).

Um exemplo dessa situação problemática é apresentado na figura 1.



Figura 1: Exemplo de janela com tamanho insuficiente
 Fonte: Certic (2007, p.5)

No contexto da Televisão Digital (TVD), há necessidade de investigar meios de implementar uma janela de Libras opcional e realizar uma análise de viabilidade de automação da geração de língua de sinais (Piccolo; Melo; Baranauskas, 2007).

O modelo de referência do sistema brasileiro (SBTVD, 2003) analisa a viabilidade e impacto da incorporação de exigências de acessibilidade, como Libras e *closed-captions*. Estes itens receberam grau de oportunidade elevado diante de cenários de diferenciação e convergência, nos quais a oferta de funcionalidades na televisão vai em direção às oferecidas na web.

Já a norma que trata da acessibilidade em comunicação na televisão (ABNT, 2005) indica o uso de uma janela para Libras, conforme a figura 2. A norma indica que a janela com a imagem do intérprete da Libras tenha as seguintes características:

- a) a altura da janela deve ser no mínimo metade da altura da tela do televisor;
- b) a largura da janela deve ocupar no mínimo a quarta parte da largura da tela do televisor;
- c) sempre que possível, o recorte deve estar localizado de modo a não ser encoberto pela tarja preta da legenda oculta;
- d) quando houver necessidade de deslocamento do recorte na tela do televisor, deve haver

continuidade na imagem da janela (NBR, 2005, p. 9).

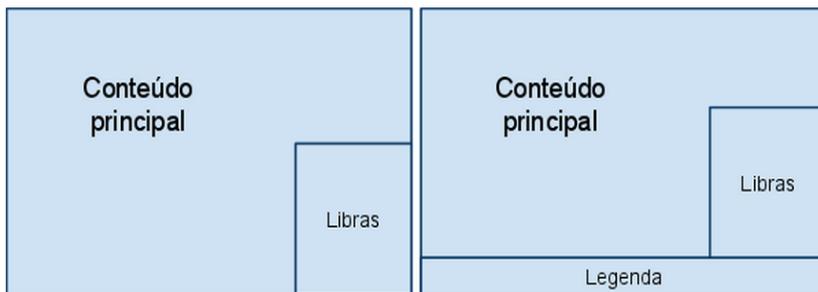


Figura 2: Espaço reservado à apresentação da Libras de acordo com recomendação de acessibilidade.

Fonte: os autores.

Na especificação de perfis para TV digital brasileira, são definidos como recursos para acessibilidade na televisão o uso de closed-captions, áudio descrição, locução, dublagem e janela de Libras (ABNT, 2007). No contexto dessa especificação, a janela de Libras corresponde ao “espaço delimitado no vídeo onde as informações são interpretadas na Libras”, e *closed-captions* consiste na “transcrição em língua portuguesa, dos diálogos, efeitos sonoros, sons do ambiente e demais informações que não podem ser percebidos ou compreendidos por pessoas com deficiência auditiva” (ABNT, 2008, p.4).

Em relação ao *closed-captions*,

A velocidade imposta pelo requisito de oferecer a transcrição literal do áudio nas legendas obriga os surdos a focarem toda sua atenção na leitura da legenda, em detrimento da visualização e interpretação das imagens, este último sendo crucial para aumentar a compreensão da mensagem televisiva (CAMBRA; LEAL; SILVESTRE, 2010, p. 49, tradução nossa).

A norma de perfis para receptores indica que a existência de recursos para tratamento do fluxo de dados referente à janela de Libras é opcional (ABNT, 2007).

Nota-se também que as especificações não estabelecem quais os mecanismos que podem ser utilizados para a codificação e envio de

conteúdo com mensagem em Libras. Também os serviços avançados de difusão de dados de TVDI não descrevem quais recursos podem ser usados na apresentação das línguas de sinais. O atendimento à acessibilidade de surdos compreende apenas o uso de legendas com caracteres textuais (ABNT, 2007).

Identifica-se a necessidade de investigar como produzir artefatos de apoio à acessibilidade para os diferentes tipos de públicos surdos (os que preferem o português e os que preferem a Libras como meio de comunicação), diante o audiovisual acessado por meio de plataformas digitais.

1.2 QUESTÃO DE PESQUISA

Diante a escassa oferta de conteúdo efetivamente acessível aos surdos, surge a seguinte questão de pesquisa: Como apoiar a produção dos artefatos para acessibilidade dos surdos ao audiovisual digital de modo a potencializar a participação dos surdos na sociedade do conhecimento?

Como problemas secundários estão questões que relacionam a produção de material audiovisual acessível com:

- Ganho de escala e redução de custos;
- Reedições e garantia dos aspectos de continuidade da interpretação;
- Adequação do intérprete a diferentes usuários e programas;
- A demanda muito grande de intérpretes em função do número de programas e quantidade de material a ser produzido;
- Apropriação dos sinais utilizados na comunicação;

Estas questões serão esclarecidas no decorrer do trabalho.

1.3 OBJETIVOS

O objetivo geral consiste na proposição de um modelo de referência que integre processos, métodos e técnicas na produção de artefatos de apoio à acessibilidade dos surdos ao conteúdo audiovisual digital.

Como objetivos específicos busca-se:

- Caracterizar os diferentes perfis de surdos;
- Identificar e caracterizar os artefatos que permitem o acesso de surdos ao audiovisual;
- Integrar os processos e mídias existentes para produção destes artefatos em um modelo de referência;
- Desenvolver uma implementação de referência na forma de um protótipo;
- Apresentar cenários de uso com produção e fruição de conteúdos com acessibilidade aos surdos.

1.4 JUSTIFICATIVA

O acesso à informação por meio das tecnologias digitais nem sempre está disponível a todos os usuários, existindo uma separação entre os que conseguem acesso à informação e seus benefícios, e aqueles que não conseguem. Oferecer acesso fácil e compreensível às plataformas tecnológicas deve extrapolar o campo econômico e fomentar a formação da cidadania por meio da inclusão digital dos diferentes grupos sociais (CASTRO, 2008).

Para a comunidade surda a demanda pela maior presença de interpretação em língua de sinais está intimamente ligada aos esforços pelo seu reconhecimento como sua língua de direito (KURZ; MIKULASEK, 2004).

A tradução do audiovisual é “um meio pragmático de integração dos deficientes auditivos com a sociedade em geral, implicando em oportunidades iguais a todos os cidadãos no acesso à informação, educação e entretenimento” (NEVES, 2005, p.309).

No Brasil, a população de surdos usuários da Libras teve a importância de sua língua de sinais reconhecida na Lei 10436 (BRASIL, 2002), onde:

É reconhecida como meio legal de comunicação e expressão a Língua Brasileira de Sinais – Libras e outros recursos de expressão a ela associados”, devendo “ser garantido, por parte do poder público em geral e empresas concessionárias de serviços públicos, formas institucionalizadas de apoiar o uso e difusão da língua Brasileira de Sinais – Libras como meio de comunicação

objetiva e de utilização corrente das comunidades surdas do Brasil (BRASIL, 2002).

O Decreto 5296 de 2004 trata do Acesso à Informação e à Comunicação e estabelece como sistemas para reprodução das mensagens veiculadas para as pessoas com deficiência o uso de legenda oculta, da janela com intérprete de Libras e da descrição e narração em voz de cenas e imagens.

Espera-se que uma maior quantidade de materiais que ofereçam vídeo em língua de sinais aumente o nível de leitura de surdos [...] em relação a sua segunda língua. Isto permite sua integração mais fácil em uma maioria social, ao mesmo tempo que preserva sua identidade, melhorando sua autoimagem e desenvolvendo sua própria cultura e linguagem (DEBEVC et al., 2009, p.282, tradução nossa).

As diferenças culturais e linguísticas dos surdos devem ser reconhecidas, de outra forma eles serão excluídos da sociedade do conhecimento (MÖBUS, 2010). Para Souza (2005), estas diferenças continuam necessitando de consideração diante das TICs, as quais são importantes nas atividades diárias e influenciam em áreas como o emprego, serviços de saúde e educação. Em relação a TV, Souza afirma que:

O descaso das emissoras de TV em relação a essa parcela de consumidores brasileiros desrespeita os preceitos da legislação do Brasil e faz com que os Surdos continuem com o desconforto de estar sempre precisando de tutores para compreender a mensagem divulgada pela televisão. Eles esperam a chegada do momento em que todos os programas de TV tenham legenda e janela com intérprete de Libras (SOUZA, 2005, p. 129).

É evidente a preferência das emissoras de TV em disponibilizar legendas em português em seus conteúdos, encontrando-se pouco conteúdo disponível com legenda de janela de Libras, porém,

A participação inclusiva dos Surdos na Sociedade da Informação deve ser efetivada de forma

autônoma e independente com condições ampliadas de acesso e uso ao ambiente informacional, assim como dos demais usuários potenciais que podem atender, independente de suas condições sensoriais e motoras (CORRADI, 2007, p.190).

Deve se atentar para a acessibilidade dos surdos, visto que a população brasileira, estimada em 190,8 milhões de habitantes pelo censo de 2010, apresenta 9,7 milhões de pessoas com algum tipo de deficiência auditiva deste, cerca de 350 mil pessoas são surdas. A tabela seguinte apresenta os estratos por tipo de surdez.

Quadro 1: Número de pessoas com dificuldades de audição
Fonte: IBGE, Censo Demográfico 2010. Tabela 1.1.

<i>Capacidade Auditiva</i>	<i>Não consegue de modo algum</i>	<i>Grande dificuldade</i>	<i>Alguma dificuldade</i>
<i>Quantidade de pessoas</i>	347 481	1 799 885	7 574 797

Do ponto de vista prático, o modelo de referência proposto identifica as possibilidades para produção ou otimização da produção de legendas em português e em Libras, possibilitando aos produtores de conteúdo um mapa para aplicação de pesquisas já existentes.

Em relação a sua relevância científica, o modelo de referência permitirá enquadrar futuras pesquisas ou projetos dentro de um esquema de trabalho que relaciona pesquisas já existentes, organizando o conhecimento existente e oferecendo uma interface de diálogo entre diferentes iniciativas.

Além disso identifica demandas para desenvolvimento de pesquisas e tecnologias associadas à oferta de audiovisual acessível aos surdos.

1.5 METODOLOGIA DA PESQUISA

Esta sessão trata da abordagem metodológica da pesquisa e da visão de mundo que orientou sua elaboração.

Quanto à natureza é pesquisa aplicada, pois visa “gerar conhecimentos para aplicações práticas, dirigidos à solução de problemas específicos” (SILVA; MENEZES, 2000, p. 20).

Quanto à abordagem do problema, é qualitativa, pois “há um vínculo indissociável entre o mundo objetivo e a subjetividade do sujeito, que não pode ser traduzida em números” (SILVA; MENEZES, 2000, p.20), sendo necessária a interpretação das observações e a atribuição de significados.

Do ponto de vista de seus objetivos, é uma pesquisa exploratória pois “visa proporcionar maior familiaridade com o problema com vistas a torná-lo explícito ou a construir hipóteses” (SILVA; MENEZES, 2000, p.20).

Do ponto de vista dos procedimentos técnicos é uma pesquisa bibliográfica e documental em conjunto com métodos de desenvolvimento de software (SILVA; MENEZES, 2000).

A pesquisa pode ser classificada como documental de fontes, ou bibliográfica, pois referência múltiplos documentos encontrados na revisão sistemática da literatura, e também documental de observadores, pois envolve mais de um examinador no plano da pesquisa por meio da aplicação do método Delphi, onde a visão de cada observador permite evidenciar tendências (POURTOIS; DESMET, 1997).

A visão de mundo que orientou a pesquisa e os procedimentos metodológicos aplicados são apresentados nas subseções seguintes.

1.5.1 A visão de mundo

As análises para a concepção do modelo de referência e sua descrição são feitas a partir de termos associados à teoria da complexidade. Esta teoria é aplicada na análise de como populações heterogêneas interagem e mantêm ou mudam suas estratégias de comunicação dependendo de como estão se saindo (SCHELLING, 1978 *apud* AXELROD e COHEN, 2000).

Os conceitos elementares da estrutura de referência para estudo de sistemas complexos proposta por Axelrod e Cohen (2000) são agente, população, estratégia, artefato, sistema, tipo, variedade, padrão de interação, espaço físico, espaço conceitual e seleção. De modo que:

Agentes, de uma variedade de tipos, usam suas estratégias, em interação padronizada, uns com os outros e com artefatos. Medidas de desempenho sobre os eventos resultantes impulsionam a seleção de agentes e/ou estratégias através de processos de cópia e recombinação propensos ao erro, mudando assim as frequências dos tipos dentro do sistema (AXELROD; COHEN, 2000, p.154, tradução nossa).

Estes conceitos são utilizados para descrever o modelo de referência. Detalhadamente, um agente consiste numa coleção de propriedades e estratégias que são usadas para a interação com artefatos e outros agentes. Os artefatos são recursos que reagem às ações dos agentes. Um conjunto de agentes com características similares constitui uma população. Dentre os agentes, os padrões de ações delimitam as estratégias. Quando tipos diferentes de agentes interagem, estes podem formar padrões de interação.

Um sistema complexo pode incluir uma ou mais populações de agentes e artefatos. Quando agentes ou artefatos compartilham características em comum, estes podem ser agrupados em um tipo. O sistema ou uma população possui uma variedade de tipos. Os agentes e artefatos podem estar distribuídos em espaços físicos ou conceituais. O espaço físico corresponde à localização temporal e geográfica, enquanto que o espaço conceitual define a afinidade que determinados agentes têm de interagir e trocar informação.

Na análise da acessibilidade do audiovisual televisivo é possível identificar dois tipos de agentes principais: os surdos e os ouvintes. Os papéis desempenhados por estes agentes podem ser o de produtor de conteúdos, que gera o material do audiovisual; os tradutores que fazem a tradução do conteúdo originalmente produzido em uma língua oral para uma língua gestual; os intérpretes que apresentam o conteúdo traduzido na língua de sinais e, finalmente, o espectador ouvinte e o próprio espectador surdo, que seleciona os artefatos que apoiarão sua estratégia de acesso ao audiovisual.

Diferentes barreiras afetam o acesso do surdo ao audiovisual, visto que este é produzido diante a norma ouvinte. A barreira sensorial é a principal delas, a qual influencia nas barreiras linguísticas. Outras barreiras afetam a acessibilidade, não somente de surdos, mas também

de outros tipos de agentes, tais como barreiras sociais, culturais econômicas, geográficas e tecnológicas.

O sistema de produção e acesso ao audiovisual pode ser analisado como um sistema complexo, pois as interações entre agentes produtores de conteúdo influenciam as ações dos agentes consumidores deste tipo de material, que por sua vez influenciam a produção dos conteúdos e também passam a participar do espaço conceitual aos quais os ouvintes tem acesso devido à inexistência de barreiras.

A interação principal dentro do sistema de produção e acesso ao audiovisual ocorre entre surdos e ouvintes, onde o audiovisual televisivo é primariamente desenvolvido para um tipo de agente e é posteriormente complementado com artefatos de apoio, para então poder ser acessado por agentes surdos.

Diversas estratégias podem ser utilizadas no acesso ao audiovisual, como a visualização da imagem das cenas, a leitura labial, a leitura de legendas escrita ou a visualização de uma janela em língua de sinais. O vídeo na janela de Libras e o *closed-captions* são exemplos de artefatos que apoiam estas estratégias de acesso.

O paradigma da complexidade permite observar os fenômenos associados à produção e oferta de conteúdo acessível como processos não lineares e recursivos, flexibilizando a abordagem de pesquisa a respeito das estratégias para produção de audiovisual acessível ao surdo.

1.5.2 Procedimentos metodológicos

As respostas à problemática e as questões problema são desenvolvidas através da visão da teoria complexidade. Os procedimentos metodológicos são apresentados ligados a cada objetivo específico.

Para o primeiro objetivo específico, “Caracterizar os diferentes perfis de surdos”, foi utilizado o método da observação direto, onde através da convivência com surdos do projeto Letras-Libras, pode-se perceber os diversos perfis deste grupo de surdos. Esta observação foi reforçada na literatura científica relacionada à problemática.

O segundo objetivo “Identificar e relacionar os artefatos que permitem o acesso de surdos ao audiovisual” foi realizado através do método de revisão sistemática da bibliografia a partir de uma chave de pesquisa que relacionava as palavras chave do problema. Assim, foram identificadas as iniciativas de pesquisa e soluções tecnológicas e metodológicas para acesso do surdo ao audiovisual, identificando os processos e papéis desempenhados na elaboração de conteúdo

audiovisual acessível. O detalhamento do procedimento de revisão sistemática é apresentado no anexo A. Após sua identificação, as iniciativas encontradas foram relacionadas por meio de taxonomia da área (ver páginas 180 e 181).

O terceiro objetivo “Integrar os processos e mídias existentes para produção destes artefatos em um modelo de referência” que apresenta os processos e artefatos passíveis de utilização em ambientes digitais foi desenvolvido baseado em cenários de uso com identificação dos itens comuns e integrados em um modelo de referência. O modelo de referência foi então validado por meio do método Delphi, apresentado na figura 3 e utilizado por Obregon (2011).

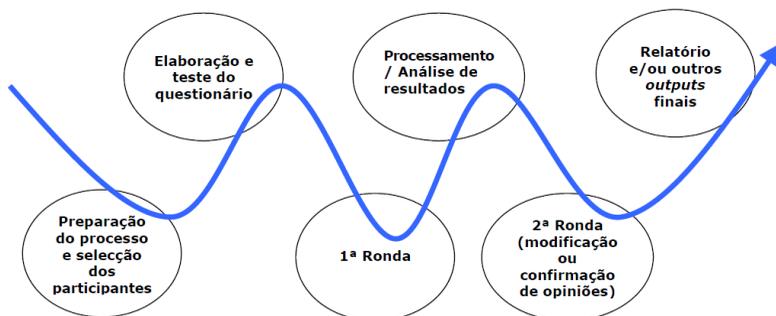


Figura 3: Estrutura base de um processo Delphi.

Fonte: Alvarenga, Carvalho e Escária (2007, p.7).

Especificamente, os especialistas consultados foram buscados na plataforma Lattes e também alguns autores dos trabalhos constantes na bibliografia utilizada. O critério de inclusão foi se o nome do pesquisador consta nos 30 primeiros resultados da plataforma Lattes para os termos de busca "surdos + legendas", "surdos+ televisão" ou "surdos + audiovisual".

O questionário inicial foi elaborado a partir de recomendações elaboradas a partir do modelo de referência.

As respostas dos especialistas ao primeiro questionário levaram a modificações no modelo de referência, que originou novo conjunto de recomendações, que foram revalidadas em um segundo questionário.

A partir da confirmação das recomendações o modelo de referência foi estabelecido e aplicado em uma implementação de referência.

O ultimo objetivo “Apresentar um cenário com produção e fruição de conteúdos com acessibilidade aos surdos”, foi desenvolvido seguindo métodos de desenvolvimento de software.

A figura 4 a seguir descreve as diferentes fases da pesquisa.

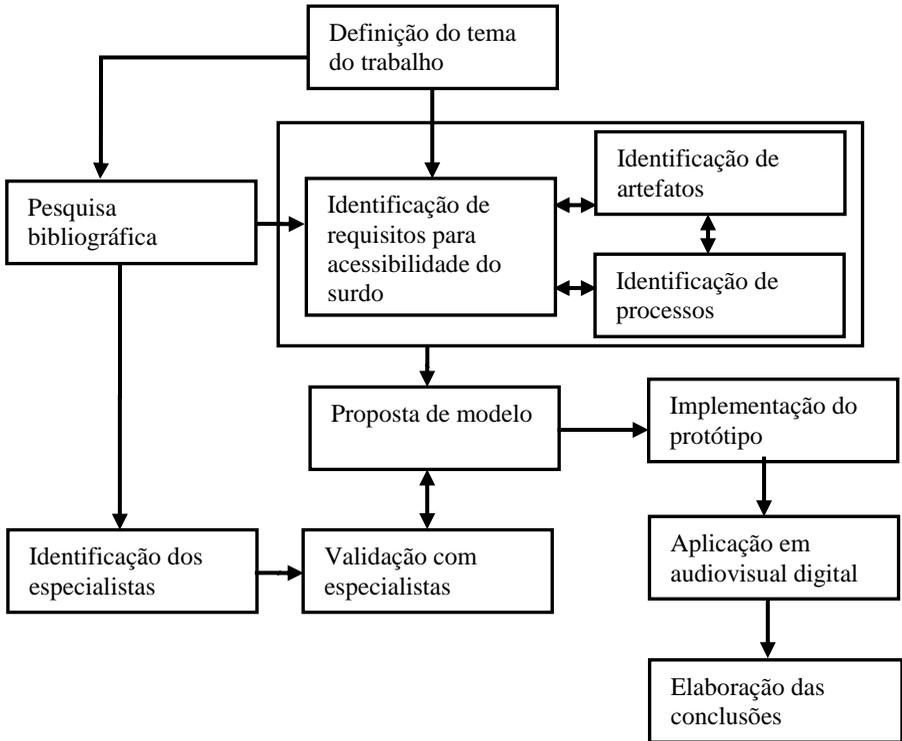


Figura 4: Fases da pesquisa.

1.6 ORIGINALIDADE

Os resultados da revisão sistemática demonstram o baixo número de trabalhos relacionados especificamente aos surdos e o audiovisual. Além disso, nos trabalhos investigados não foram observadas iniciativas que estabelecem meios de apoio às relações entre surdos e audiovisual por meio de diferentes mídias. Alguns trabalhos englobam apenas aspectos relacionados às legendas textuais, enquanto que outros consideram apenas o uso das línguas de sinais. A inovação e ineditismo está na abordagem que visa integrar as duas línguas aplicadas ao audiovisual, com destaque para a aplicação da escrita de sinais na legendagem.

Pretende-se relacionar essas iniciativas em um documento que oriente o projeto e desenvolvimento de sistemas de apoio à produção de audiovisual acessível aos surdos.

1.7 ESCOPO E LIMITES DA PESQUISA

A pesquisa oferece recursos que permitem ampliar a acessibilidade dos surdos ao audiovisual, potencializando sua participação na sociedade do conhecimento.

O foco está na identificação e integração de processos e mídias utilizados no desenvolvimento e oferta de audiovisual digital acessível ao surdo, não contemplando questões relativas à material audiovisual em meios analógicos.

Devido à sua natureza interdisciplinar, a pesquisa referencia trabalhos da área de linguística aplicada e tradução, de produção de audiovisual e de técnicas de visão computacional e animação, entretanto não engloba a análise crítica e aprofundada de procedimentos como as estratégias de tradução, apenas os referencia no processo de produção do audiovisual.

O conteúdo audiovisual abordado neste trabalho não contempla os materiais produzidos 'ao vivo'.

Considera-se que os materiais que serão tornados acessíveis já foram produzidos. Esta limitação decorre do fato que se a acessibilidade também pode ser considerada no momento da produção do audiovisual, evitando-se, por exemplo, o uso de linguagem complicada, ou então em cenas que já são gravadas com o intérprete.

Os processos associados à tradução demandam tempo para o intérprete ou tradutor realizarem, entretanto os cenários com produção

de conteúdo ao vivo podem ser contemplados com o uso de técnicas de tradução automática, mas este não é o foco do trabalho. A enumeração de técnicas de tradução automática visa apenas reduzir a carga de trabalho dos tradutores manuais.

A indicação da possibilidade de uso das técnicas de tradução automática não atende completamente os requisitos para acesso do surdo ao audiovisual, visto que estas técnicas ainda demandam aprimoramento.

1.8 ADERÊNCIA AO PROGRAMA

O programa de Engenharia e Gestão do conhecimento (EGC) possui três áreas de pesquisa principais voltadas à disseminação do conhecimento: a área de Engenharia do conhecimento; a Gestão do Conhecimento; e a Mídias do Conhecimento.

As mídias do conhecimento (*knowledge media*) permitem a interação e a comunicação por meio de artefatos que incorporam dados e processos voltados ao conhecimento. Elas representam espaços de realização de tarefas em que as pessoas desenvolvem seus trabalhos e também são espaços interpessoais, onde a comunicação ocorre. O objetivo central das mídias do conhecimento é ajudar comunidades de indivíduos a pensar, comunicar, aprender e criar conhecimento (Baecker, 1997).

A mídia do conhecimento pode ser aplicada em diferentes contextos, como na mudança de uma rede fechada de especialistas para uma comunidade que interage num espaço de troca de conhecimentos que engloba diferentes tipos de conhecimentos externalizados. Sua presença também pode ser constatada no desenvolvimento de linguagens comuns em comunidades definidas, e na tradução entre as linguagens destas comunidades (STANOEVSKA-SLABEVA, 2002).

A mídia do conhecimento deve oferecer então os recursos que permitam a disseminação do conhecimento, nesse contexto, busca-se uma mídia adequada para a produção e publicação de conteúdo audiovisual digital para surdos.

A gestão do conhecimento busca identificar as estruturas de apoio aos processos de uso e disseminação do conhecimento por meio das línguas de sinais e legendas textuais.

A engenharia do conhecimento oferece subsídios para a tradução e transcrição automática da língua de sinais, assim como a simplificação automatizada de textos, o que permite aperfeiçoar o processo de desenvolvimento de conteúdos acessíveis.

No contexto do programa de Engenharia e Gestão do Conhecimento, esta pesquisa enquadra-se na linha de pesquisa de Teoria e prática em Mídia e Conhecimento, pois estuda sistemas multimídia para a disseminação do conhecimento.

1.9 RESULTADOS ESPERADOS

Como resultado busca-se a definição de um modelo de referência que identifique artefatos, papéis e processos e oriente sua aplicação em cenários de produção de artefatos de acessibilidade do surdo ao audiovisual.

Um modelo de referência consiste em um

“framework abstrato para o entendimento de relacionamentos significantes entre entidades de determinado ambiente, e para o desenvolvimento de padrões ou especificações que apoiam este ambiente. Um modelo de referência é baseado em um número reduzido de conceitos unificadores e pode ser utilizado como base para a educação e explicação destes padrões para não especialistas. Um modelo de referência não está diretamente relacionado a determinado padrão, tecnologia ou outro detalhe de implementação concreta, mas visa oferecer uma semântica comum que pode ser usada de forma não ambígua entre diferentes implementações.” (OASIS, 2011, sem paginação, tradução nossa).

O modelo de referência permite a idealização de sistemas para a produção de conteúdos audiovisuais acessíveis, estabelecendo um vocabulário comum para a descrição de métodos e técnicas para produção de artefatos de apoio à acessibilidade dos surdos.

1.10 ESTRUTURA DA TESE

Esta tese está estruturada em sete capítulos. O segundo e próximo capítulo fundamenta a variedade de condições que influenciam a formação dos surdos, apresenta aspectos relacionados à acessibilidade dos surdos ao conteúdo audiovisual diante às características das mídias

disponíveis na TVD e da web e caracteriza os processos de tradução relacionados ao audiovisual.

O terceiro capítulo identificando os artefatos para acessibilidade e os processos e técnicas para sua produção.

O quarto capítulo descreve o modelo de referência voltado à escolha e integração de métodos e alternativas tecnológicas para o desenvolvimento de artefatos de acessibilidade.

O quinto capítulo descreve a aplicação do modelo de referência em uma situação com produção de conteúdo acessível aos surdos destinado a duas plataformas digitais: uma para cenários de televisão digital e outra na web.

O sexto capítulo apresenta os resultados obtidos, as conclusões e perspectivas de trabalhos futuros que foram identificadas no desenvolvimento da tese.

1.11 ENVOLVIMENTO PESSOAL COM A PESQUISA

Nesta sessão busca-se apresentar fatores que influenciaram na escolha deste tema de pesquisa e na forma com que as soluções para acessibilidade são abordadas.

O fato principal para a escolha do tema foi a participação deste autor na implantação do curso de licenciatura e bacharelado em Letras Libras na Universidade Federal de Santa Catarina que iniciou no ano de 2006. Este curso demandou a configuração e adaptação de um Ambiente Virtual de Aprendizagem e o desenvolvimento de conteúdos em hipermídia direcionados ao público surdo.

No processo de implementação dos sistemas e materiais educacionais foram encontrados desafios interessantes de serem superados e também a relevância das mídias digitais no apoio à comunicação dos surdos.

Em 2009 participou-se do grupo de pesquisa de formação de recursos humanos para TV digital, que investigou meios para o desenvolvimento de conteúdos para esta plataforma.

Estas vivências, além do interesse por modelagem digital e programação, levaram à exploração das possibilidades de aprimoramento dos processos de produção de conteúdo baseado em mídia digital, especificamente os meios para oferta de conteúdo acessível ao público surdo.

1.12 RESUMO DO CAPÍTULO

Este capítulo inicial apresentou a problemática da pesquisa e a metodologia utilizada na busca de uma solução.

A questão de pesquisa busca responder como oferecer acessibilidade ao conteúdo audiovisual aos surdos, em termos de processos e artefatos e considerando os diferentes aspectos do público surdo.

O capítulo seguinte aborda a questão dos públicos surdos com mais profundidade.

2. ACESSIBILIDADE AO AUDIOVISUAL: OS SURDOS, A TRADUÇÃO E AMBIENTES DE ACESSO

Para tratar da acessibilidade dos surdos ao audiovisual foi considerado necessário:

- Esclarecer quem são os sujeitos surdos, de modo a compreender suas diferenças;
- Apresentar as características das plataformas digitais onde o acesso ao audiovisual pode ocorrer, de modo a identificar as possibilidades e limitações de cada ambiente;
- Analisar os conceitos relacionados à tradução de conteúdo audiovisual, de modo a entender e apoiar os processos que ocorrem na oferta de acessibilidade;

As próximas sessões tratam respectivamente destes temas.

2.1 O PÚBLICO SURDO

A acessibilidade do audiovisual pode ser promovido pela oferta da transcrição textual daquilo que é falado e de marcações que indicam eventos sonoros, como campainhas e barulhos. Entretanto esta concepção de garantia de acesso não contempla adequadamente as necessidades dos sujeitos surdos, visto que existem em um contínuo entre o extremo dominado por uma cultura ouvinte e outro extremo, identificado pela cultura surda, conforme afirma Mello (2001):

Enquanto que surdos usuário de Línguas de Sinais “estão mais próximos de uma ‘massificação’ da cultura surda, que tem na língua de sinais a sua manifestação maior de cultura; os oralizados se aproximam mais das manifestações da cultura ouvinte, onde privilegia-se a habilidade da fala e eficácia em leitura labial (MELLO, 2001, sem paginação).

Perante esta visão da existência de um contínuo, o próximo tópico explora a variedade de tipos dentro da população surda.

2.1.1 Uma diversidade de visões

Toda classificação é uma forma de redução da complexidade relacionada à diversidade proveniente da combinação dos diferentes fatores que influenciam na constituição do sujeito surdo (TORREZ; MAZZONI; MELLO, 2007). Um filho surdo pode nascer em famílias de pais ouvintes, um filho ouvinte pode nascer em uma família de pais surdos, ou apenas sua mãe pode ser surda. O filho pode ficar surdo em uma família de ouvintes, sem contato anterior com questões relacionadas à surdez. A surdez também pode surgir antes ou depois da formação das estruturas de linguagem (SILVA; PEREIRA; ZANOLLI, 2007).

apesar de todas caracterizações apresentadas sobre cada grupo específico de surdos[..] há sempre o perigo imposto pela armadilha da generalização. Existem casos raros, que podem ser isolados ou pertinentes a certos grupos, no que se refere à forma de aceção do mundo pelo surdo. Havendo uma enorme variação de situações em que o surdo

pode ser bom em leitura labial, mas ruim na escrita, bom em sinais e bom na escrita, é muito difícil ainda, devido à escassez de pesquisas, estabelecer uma base mais resistente a respeito de cada caso específico. Portanto, aqui fica registrado o alerta a respeito das generalizações apresentadas (MELLO, 2001, sem paginação).

Para fins estatísticos, o IBGE (2010) define “deficiência auditiva” como a incapacidade auditiva que o sujeito encontra mesmo com o uso de aparelho auditivo, levando em conta a percepção do sujeito sobre sua audição. Esta podendo ser dividida em:

- Incapaz de ouvir (pessoa se declara totalmente surda);
- Grande dificuldade permanente de ouvir (pessoa declara ter grande dificuldade permanente de ouvir, ainda que usando aparelho auditivo);
- Alguma dificuldade permanente de ouvir (pessoa declara ter alguma dificuldade permanente de ouvir, ainda que usando aparelho auditivo) (IBGE, 2010, sem paginação).

Além dessa classificação, a caracterização da deficiência auditiva pode ser realizada tanto “em função do local de acometimento, quanto ao grau de privação sonora, em relação ao desenvolvimento da fala e pode ser classificada ainda como congênita ou adquirida” (SANTOS, 2008, p.4).

Existem diferentes visões também relativas à educação e formação dos sujeitos surdos, sua cultura e sua interação com o espaço conceitual formado pelo mundo dos ouvintes. Historicamente, o oralismo e a comunicação total podem ser apresentados como abordagens para compreensão do sujeito surdo e sua comunicação, porém estas abordagens não respeitam adequadamente a constituição sociocultural do surdo.

A cultura oralista, por exemplo, desconsidera a existência de uma cultura surda e afirma que o surdo, para sua inclusão na sociedade, deve ser fluente na língua oral e que sua expressão deve utilizar recursos como a verbalização, leitura labial e escrita.

Fazer uso da linguagem escrita significa, para os surdos, apropriar-se de um conhecimento social e cultural de grande amplitude. Essa apropriação se dá pelo diálogo entre diferentes linguagens sociais, diferentes discursos, na pluridiscursividade de enunciados e na multiplicidade de vozes. Isso pode tornar o surdo integrante da nação a que pertence, na medida em que, ao se apropriar da linguagem escrita em língua portuguesa, ele pode estabelecer e ampliar suas interações, negociar sentidos, trocar conhecimentos, inserir-se no mundo discursivo e produzir seus próprios enunciados (ARCOVERDE, 2006, p.257).

Um elemento representativo da cultura oralista é o implante coclear, que consiste num dispositivo eletrônico que pode ajudar a proporcionar algum sentido de audição de sons a alguém que seja profundamente surdo. O dispositivo consiste em uma parte externa que fica atrás da orelha e uma parte interna colocada cirurgicamente sob a pele. Em relação a este tipo de implante, existe um benefício substancial na fala e no vocabulário quando a criança surda recebe seu implante antes da idade de dois anos e seis meses. Os benefícios aumentam com o tempo de uso e diminuem sistematicamente em relação à idade em que o implante foi realizado (HAYES; KESSLER; TREIMA, 2011). Porém a visão dos surdos sobre este procedimento e a inserção dos surdos na sociedade por meio da oralização é divergente, conforme afirma Wrigley (1996).

O ouvinte há de pensar que isto [a surdez] tem algo a ver com a ideia de ter estado doente, ou com o sentimento de perda, ou senso de culpa, pois, para o ouvinte, a surdez representa perda de comunicação, exclusão, banimento, solidão, isolamento. Para os surdos a explicação é totalmente diferente: alegar uma surdez de nascença significa não estar contaminado pelo mundo dos ouvintes e suas limitações epistemológicas de som sequencial (WRIGLEY, 1996, p.39).

A oralização do surdo é uma abordagem que visa possibilitar ao surdo a comunicação com seus pares ouvintes, entretanto esta

abordagem não pode ser tida como a única alternativa para o desenvolvimento do sujeito surdo (SILVA, 2001).

Médicos, fonoaudiólogos, professores, familiares, muitos querem “desensurdecer” os surdos, anular sua existência como sujeitos surdos e junto com eles, anular suas línguas, sejam elas em suas formas sinalizadas ou escritas (BARROS, 2008, p. 17).

A abordagem da comunicação total, por sua vez, consiste numa estratégia que busca a integração do surdo por meio do uso de gestos e sinais, sendo fortemente influenciada pela cultura oralista. Ela visa desenvolver os processos comunicativos entre surdos e entre surdos e ouvintes, fomentando tanto os aspectos cognitivos quanto os linguísticos. Para isso utiliza todos os recursos que facilitem a comunicação, como o alfabeto datilológico, os sinais da língua de sinais, gestos e expressão corporal (GOLDFELD, 1997).

Na comunicação total o surdo está exposto a um conjunto de possibilidades comunicacionais, na perspectiva de facilitar a sua expressão. Contudo, a comunicação total não lhe propicia o ambiente linguístico adequado a um desenvolvimento da língua materna que permita a estruturação do pensamento. Esses contextos comunicacionais, por não oportunizarem o acesso à língua de sinais, não podem oferecer informações mais complexas e geralmente acabam restringindo-se a conversas mais concretas, em que as manifestações dos surdos tendem a ficar num nível mais elementar (SILVA, 2001, p. 53).

A comunicação total valoriza o papel da família no processo educativo da criança surda, definindo o surdo não como portador de uma deficiência, mas como sujeito que possui uma diferença sensorial que repercute em suas relações sociais e em seu desenvolvimento afetivo e cognitivo (SILVA, 2001).

Silva et al. (2007) ao estudarem famílias com filhos surdos constata que uma família de pais ouvintes optam pela educação de um filho surdo segundo a cultura de oralização, e encaram a surdez como

uma deficiência. Já filhos surdos de pais surdos tendem a ter uma educação bilíngue onde a surdez é encarada como uma mera diferença.

Existem diferenças entre as estratégias de leitura de ouvintes e de surdos com perda auditiva pré-linguística: surdos pré-linguísticos desenvolvem estratégias de leitura que focam primariamente em propriedades visuais ao processarem as palavras escritas, enquanto que os ouvintes baseiam-se nos constituintes fonológicos das palavras (MILLER, 2009).

Assim surge uma barreira para a leitura de textos por parte dos surdos pré-linguísticos, não originada a partir do déficit de leitura no nível léxico, mas sim pela eficiência no reconhecimento dos significados finais das palavras a partir de sua estrutura dentro do contexto de uma sentença, de modo que, apesar do uso de estratégias similares de leitura, ainda persiste a diferença do nível de leitura entre surdos e ouvintes (MILLER, 2009).

A respeito das diferentes fases linguísticas de crianças surdas, Lichtig, Couto e Leme (2008), em uma visão pragmática, constatam que

Pesquisas na área da surdez mostram que o desenvolvimento da forma e conteúdo da língua oral nas crianças surdas é mais lento do que em seus pares ouvintes. Em média, cinco meses do desenvolvimento da oralidade em uma criança ouvinte equivalem a um ano do desenvolvimento da criança surda, crianças estas consideradas como bem sucedidas em programas de intervenção fonoaudiológica (LICHTIG; COUTO; LEME, 2008, p.251).

Ao se analisar a dinâmica das famílias de surdos, Brito e Dessen (1999) afirmam que

faz-se necessário conhecer a dinâmica de funcionamento dessas famílias no período inicial após o nascimento da criança, ou após o evento que a tornou surda, e durante todo o processo de adaptação à nova situação, focalizando como evoluem as interações e relações entre a criança surda e todos os membros de sua família (Brito; Dessen, 1999, p.444).

Nesse sentido, Araujo, 2008 justifica que

Ao contrário do que se pensa, os surdos têm maior dificuldade de leitura na língua oral de seu país, porque esta funciona como segunda língua para eles, sendo a língua de sinais sua primeira língua. Essa visão bilíngue vem sendo absorvida vagarosamente, visto que o bilinguismo só começou a ser discutido a partir de 2002 com o reconhecimento da Língua Brasileira de Sinais [...] como a língua da comunidade surda do nosso país [...]. Na verdade, os surdos brasileiros ainda aprendem o português como língua materna, apresentando, por essa razão, baixa proficiência no idioma, principalmente no que diz respeito à leitura e escrita (ARAUJO, 2008, p.63).

É importante considerar que a formação inicial do sujeito surdo deve se basear em sua língua materna, aquela que está apta a perceber, para então partir para sua inclusão social. Nessa perspectiva, a comunicação total não considera adequadamente os passos iniciais da constituição do sujeito. Deste modo, a segunda língua, no caso o português, não pode ser aprendida se o surdo não tiver uma língua materna que já lhe tenha propiciado o desenvolvimento do pensamento (BAKHTIN, 1990).

O próximo tópico descreve a abordagem pela educação bilíngue, que considera a real natureza do sujeito surdo.

2.1.2 A Educação Bilíngue

O bilinguismo considera a questão surda como de natureza linguística e cultural, onde o uso de uma língua gera uma cultura que reflete a realidade de um grupo de sujeitos (FLEURY, 2010).

As comunidades surdas defendem a educação bilíngue, onde se utilizam as línguas de sinais não com o objetivo da oralização, mas para permitir aos surdos participarem dos debates da sociedade atual no mesmo nível de igualdade dos ouvintes e ao mesmo tempo reconhecendo as especificidades de sua língua visual-espacial, permitindo que esta sirva de instrumento para a identidade cultural dos surdos (SKLIAR, 1997).

A figura 4 apresenta um chamado à comunidade surda nacional para defesa de um modelo educativo que opte pelo bilinguismo.



Figura 4: Manifestação em prol da educação bilíngue

Na abordagem bilíngue, o surdo é considerado um ser social e cultural que interage tanto com sujeitos surdos quanto com sujeitos ouvintes. Para tal deve ser alfabetizado em língua de sinais e também em língua oral escrita, no caso brasileiro, na Libras e no português.

O surdo ao ser encarado com um sujeito bilíngue, utiliza a língua de sinais como sua língua natural, para afirmação de sua identidade e a língua oral para interação com a sociedade ouvinte.

A figura 5 apresenta material didático utilizado em uma escola de educação especial pela professora Sonia Messerschmidt para ensino de escrita de sinais e português.



Figura 5: Material de ensino de línguas de sinais na Escola Estadual de Educação Especial Dr. Reinaldo Coser

Fonte: Sonia Messerschmidt (<http://www.facebook.com/sonialibras2011>)

Outro exemplo de material para educação bilíngue é apresentado por Flood (2002), na figura 6, ao estudar como crianças se apropriam dos sistemas de escrita da língua de sinais.



Figura 6: Exemplo de material didático para alfabetização de surdos
Fonte: Flood (2002, p.140).

Os sistemas de escrita de sinais são apresentados juntamente com as Línguas de Sinais (LS), língua de natureza gestual e visual, que é um importante modo de comunicação dos surdos e de formação de sua cultura.

2.1.3 As Línguas de Sinais

Estudos sobre as Línguas de Sinais (LS) “provaram que ela tinha um valor linguístico semelhante às línguas orais e que cumpria as mesmas funções, permitindo ao surdo se expressar em qualquer nível de abstração” (SILVA, 2001, p.48).

As comunidades surdas estão habituadas a utilizar línguas de sinais entre si. Por muitos anos, a línguas de sinais não foram consideradas línguas legítimas. Elas eram erroneamente consideradas simples mímica que espelhava a comunicação oral e muitos pensavam que esta era universalmente entendida (NEVES, 2005, p.93).

Nas línguas orais as palavras são produzidas pelo trato vocal e são percebidas como sons sequenciais, já nas línguas de sinais as palavras podem ser produzidas simultaneamente, pela articulação e localização da mão, movimentos manuais, expressões faciais, movimentos da cabeça, postura corporal, sendo percebidas visualmente e em paralelo, explorando ao máximo as relações temporais e espaciais de modo a estabelecer um significado (CAPLIER et al., 2007).

Ao utilizar a língua de sinais o indivíduo surdo coloca o rosto do outro no centro deste campo, acompanhando os gestos manuais com a periferia do campo visual. Portanto, esta maior habilidade para discriminar e seguir estímulos, no caso o movimento das mãos, na periferia do campo visual é uma grande vantagem para o surdo, a qual não é aproveitada caso ele não utilize a língua de sinais (RODRIGUES, 1993).

As línguas de sinais são um meio de comunicação que possuem estruturas linguísticas próprias (BRITO, 1995; QUADROS; KARNOPP, 2004), sendo exemplos de línguas de sinais a *American Sign Language* (ASL), a *Langue des Signes Française* (LSF) e a Língua Brasileira de Sinais (Libras).

No contexto brasileiro, a Libras não constitui uma representação alternativa do português, de modo que os termos de uma sentença da língua falada não possui correspondência direta com os termos da língua de sinais, e vice-versa. A simples realização de sinais traduzindo

diretamente os termos da língua falada constitui o Português Sinalizado, que não permite a compreensão do conteúdo pelo surdo e impede a comunicação efetiva.

A questão língua de sinais *versus* português sinalizado é demonstrado por Silva (2001) por meio da frase “No domingo passado eu jantei no restaurante”. Na Libras a frase pode ser estruturada na seguinte forma²:

“DOMINGO / PASSADO / EU IR / RESTAURANTE / COMER-NOITE”.

Já no português sinalizado, a estrutura da frase incluiria duas vezes a sílaba “no”, no início e antes da palavra “COMER/NOITE”, excluindo-se o verbo direcional “IR” e substituindo a palavra “COMER-NOITE” por “jantei e restaurante”.

Para a comunidade surda, a LS é sua língua de escolha, de modo que para o surdo, “a leitura de texto em uma língua falada é semelhante ao uso de uma língua estrangeira” (KENNAWAY; GLAUERT; ZWITSERLOOD, 2007, p.16, tradução nossa).

Nessa perspectiva a formação do sujeito deve considerar o uso das línguas de sinais, que além de meio de comunicação na sociedade configura também um meio de formação de identidade e cultura.

Um dos aspectos no campo linguístico é que a língua de sinais, enquanto experiência visual, passa a ser um direito de luta que articula a organização política dos movimentos surdos e ‘não um instrumento imediato a serviço da língua majoritária’ (SILVA, 2001, p.19).

Em relação à sua gramática, as LS podem ser analisadas quanto à sua fonologia, morfologia e sintaxe.

A fonologia, no contexto das línguas de sinais, define um sistema com parâmetros que formarão os sinais, identificando parâmetros de configuração de mão, pontos de articulação, movimentos, expressão facial e corporal e a orientação/direção, conforme afirma Souza (2005b):

² Neste exemplo foram usadas glosas com palavras do português para representar os sinais da Libras.

A fonologia das línguas de sinais estuda as configurações das mãos, o ponto de articulação e os movimentos, que são tidos como parâmetros primários. A região de contato, orientação e disposição das mãos são considerados parâmetros secundários (SOUZA, 2005b, p.24).

Os parâmetros fonológicos da língua de sinais podem ser combinados para criar sinais significativos lexicamente (EMMOREY; BELLUGI; KLIMA, 1993).

Já a morfologia dos sinais preocupa-se em descrever características como gênero, grau, indicação de tempo, negação e a datilologia nas LS. A datilologia é utilizada na pronúncia de nomes ou quando não existe o sinal correspondente na LS, onde a palavra é soletrada utilizando sinais manuais que remetem ao alfabeto da língua oral.

Por sua vez, a sintaxe trata das inter-relações entre os elementos de uma frase e a combinação de sentenças, o estudo dos verbos e pronomes pessoais por meio de sinais.

A morfologia e a sintaxe das línguas de sinais determinam a estrutura interna das palavras e das frases que reflete o sistema computacional da linguagem (QUADROS, 2004, p.22).

Os parâmetros fonológicos são descritos por sistemas de notação utilizados no registro escrito das LS. Estes sistemas para escrita em LS são descritos no próximo tópico.

2.1.4 Escrita das Línguas de Sinais

As línguas de sinais não tem tradição de registro escrito, porém destaca-se a necessidade do uso de um sistema de escrita iconográfica, que preserve a característica viso espacial das línguas de sinais (CORREIA; LIMA; LIMA, 2010).

De modo a preservar as características visual e espacial das LS no registro escrito, a escrita de sinais adota “um sistema de representação gráfica das línguas de sinais que permite através de símbolos visuais representar as configurações das mãos, seus movimentos, as expressões faciais e os deslocamentos corporais” (STUMPF, 2000, sem paginação).

A Escrita da Língua de Sinais ocupa um ‘lugar’ de marcador cultural, de tradução cultural surdo, pois retrata a diferença e experiência de ser surdo no sentido mesmo de disseminador de uma cultura, que se alicerça em conceitos como a diferença e de experiência visual (ZAPPE, 2010, p.63).

Um sistema de notação escrita de sinais não representa seus significados diretamente, mas descreve apenas seus aspectos físicos (KENNAWAY, 2004). Existindo diferentes sistemas de escrita para as línguas de sinais, por exemplo, o *Stokoe*, o *HamNoSys* e o *SignWriting* (MCCARTY, 2004). O sistema de escrita *Stokoe* foi o primeiro sistema para escrita de sinais, tendo um objetivo mais científico do que prático em relação a sua utilização pelos surdos. O *HamNoSys*, assim como o *Stokoe*, foi projetado com objetivos mais científicos. Já o *SignWriting* visa a comunicação baseada nos sinais. Estas três formas de registro das línguas de sinais, aplicados à *American Sign Language* (ASL) estão apresentados na figura 7 abaixo.

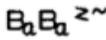
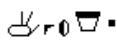
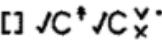
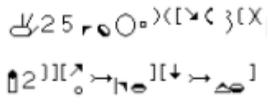
	Sign Writing	Stokoe	HamNoSys
what?			
quote			
three			
bears			
Goldilocks			

Figura 7: Três notações para o registro escrito das línguas de sinais.
 Fonte: Maurer, Stubenrauch e Camhy (2003, p. 334).

Além destes sistemas de escrita, o sistema de escrita *Elis*, apresentado por Barros (2008), consiste em um sistema de escrita linear, assim como o *Stokoe* e *HamNoSys*. Barros (2008) descreve sua aprendizagem por um grupo de surdos, os quais propõem a produção de dicionários e materiais didáticos utilizando este sistema de escrita. As figuras 8 e 9 apresentam respectivamente um texto em Libras escrito com o *Elis* e sua tradução para o português.

-I.° ΠUθ',
 .I.ΠΦΟ' // -ΠΦ' Δ' Δ -ΠΦ@°ΔΔΦθ <Γ' ΔΔ' .I.vΦ#'
 // <ΓΔΦ#° -I.° μ#Δ', .I.ΠΟ' ΠΔΔ', <Γ' Δ#° .Δ.ΔΔΦ#° <°
 <ΓΠΣΔ .I.ΠΦΟ' // <ΓΔΦ#° -I.vΦ → ΔΓω° // -ΔvΦ0'°
 -I.vΔ' Δ!

Figura 8: Texto no sistema de escrita Elis.

Fonte: Barros (2008, p.137).

Mariângela,

Nós queremos te agradecer, pois estamos felizes com o curso de janeiro, em que você nos ensinou a ELiS. Vamos chorar. Nós aprendemos coisas diferentes.

Será bom e importante ensinar as crianças para que no futuro possam se desenvolver.

Obrigada!

Figura 9: Tradução livre para o português do texto em LS.

Fonte: Barros (2008, p.139).

Em relação ao sistema de escrita *HamNoSys*, ilustrado na figura 10, Kennaway (2004) afirma que este sistema apresenta problemas referentes a omissão de informações, dificuldade de representação quanto a direção do dedo estendido, ambiguidade da fonética gestual e escopo.

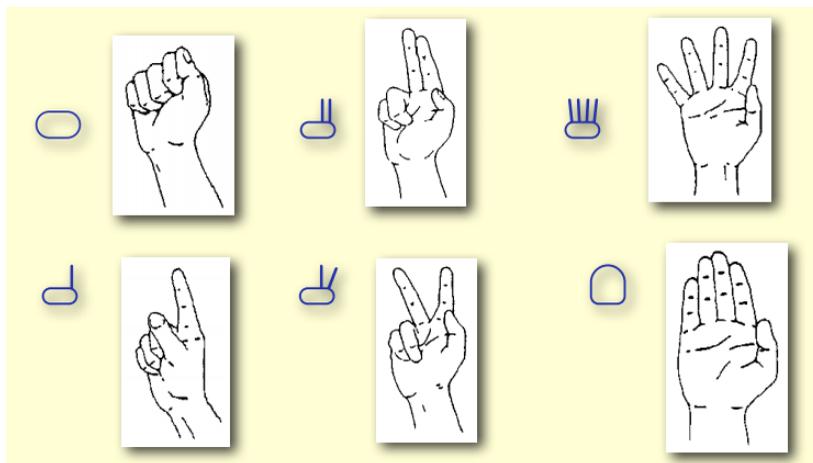


Figura 10: exemplos de configurações de mão e sua notação correspondente em HamNoSys

Fonte: Hanke (2009, p.6)

A escrita de sinais *SignWriting* consiste em representações gráficas da língua de sinais (SUTTON, 1990). Exemplos de escrita de sinais em *SignWriting* e seus sinais correspondentes podem ser vistos na figura 11.

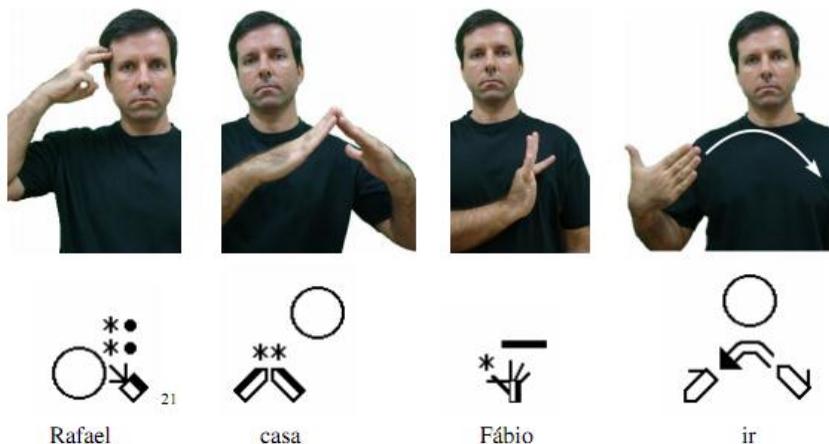


Figura 11: exemplo de correspondência entre sinais e escrita de sinais.

Fonte: Silva (2009, p.66)

Deve se notar que a escrita de sinais ocorre do ponto de vista do sinalizante. A figura 12 apresenta um exemplo específico de correspondência entre configuração de mão, vista sob essa perspectiva, e um caractere em escrita de sinais.

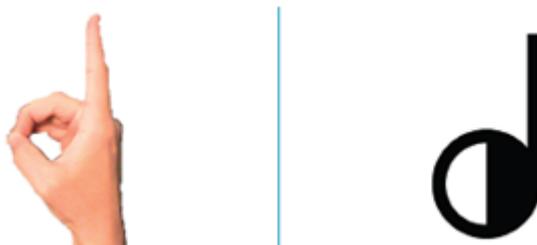


Figura 12: Correspondência entre configuração de mão e seu símbolo em escrita de sinais.

Fonte: Sutton e Frost (2011)

Exemplos de expressões faciais descritas com escrita de sinais são apresentados na figura 13, apresenta a correspondência entre símbolos referentes à bochechas.

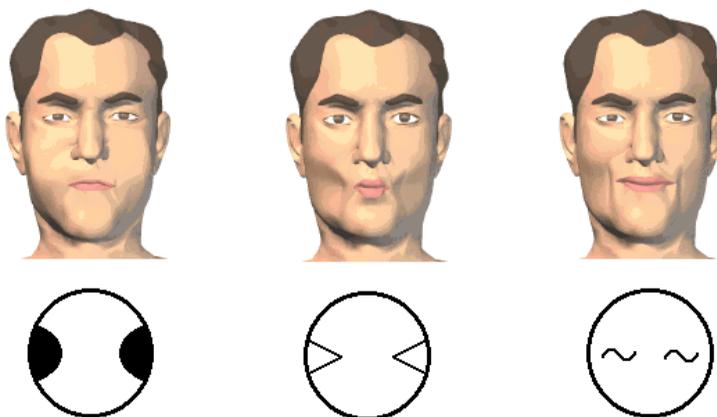


Figura 13: correspondência de símbolos com expressões faciais referentes às bochechas.

Fonte: Parkhurst e Parkhurst (2010, p.143)

Na escrita de sinais *SignWriting* (SW) um sinal da linguagem, correspondente a um significado que é transcrito em um conjunto de

símbolos posicionados em um canvas 2D chamado *SignBox* (SUTTON, 2008).

A escrita de sinais *SignWriting* pode ocorrer de três modos (WÖHRMANN, 2003):

- escrita com o corpo inteiro, que “utiliza a figura completa do corpo, uma forma mais fácil de ser entendida pelos iniciantes”;
- a escrita de sinais padrão, que “utiliza a figura com símbolos tornando o sinal uma unidade visual. É a forma considerada padrão no uso da escrita da língua de sinais”;
- e a escrita simplificada, que “é uma forma simplificada da escrita padrão que exclui alguns símbolos de maneira a facilitar a redação escrita a mão”.

Existem dois tipos de sequenciamento em relação aos sinais no *SignWriting*: a sequência de soletragem do sinal, e a sequência de símbolos do sinal (SUTTON, 2008).

A *sequência de soletragem* do sinal define a ordem de leitura dos símbolos dentro de um sinal, dividindo ele em sílabas que estabelecem qual grupo de símbolos aparecem primeiro, por segundo, e assim sucessivamente. Um sinal escrito não é lido diretamente da direita para a esquerda, mas sim de acordo com a sequência de soletragem ou pronuncia, de modo que a pronuncia de sinais tem aspectos sequenciais e simultâneos.

Na escrita de sinais *SignWriting*,

Os símbolos são unidades visuais relacionadas ao centro do sinal, onde o leitor busca o formato de mão que inicia o sinal e segue os símbolos de movimento para a segunda posição da mão. Este processo, apesar de não ser linear, cria a sequência de soletragem dos sinais (SUTTON, 2008, p.3).

Especificamente, cada sinal é composto por 4 sílabas, de tal modo que os símbolos que ocorrem em cada sílaba correspondem a movimentos realizados ao mesmo tempo. As quatro sílabas consistem em:

Sílaba 1: Mãos

1. Posição inicial da mão dominante
2. Posição inicial da mão não dominante

Sílaba 2: Movimento

1. Primeiro movimento da mão dominante
2. Primeiro movimento da mão não dominante
3. Segundo movimento da mão dominante
4. Segundo movimento da mão não dominante
5. Dinâmica da mão dominante;
6. Dinâmica da mão não dominante

Sílaba 3: Mãos

1. Posição final da mão dominante
2. Posição final da mão não dominante

Sílaba 4: Ordenamento detalhado

1. Localização da mão dominante
2. Localização da mão não dominante
3. Cabeça
4. Face
5. Pescoço
6. Ombro
7. Torso
8. Braços & Pernas

A primeira sílaba corresponde à configuração das mãos, a segunda sílaba corresponde ao movimento realizado, a terceira sílaba indica a posição/configuração final das mãos e a quarta sílaba indica localização, expressões faciais, aspetos da cabeça e do corpo. A sequência padrão é mão dominante, mão não dominante (sílabas um e três); primeiro movimento dominante, primeiro movimento não dominante, segundo movimento dominante, segundo movimento não dominante, dinâmica dominante, dinâmica não dominante; localização da mão dominante, localização da mão não dominante, cabeça, rosto, pescoço, ombro, torso, braços e pernas (SUTTON, 2008a).

A figura 14 exemplifica a aplicação da ordem de soletração do sinal.

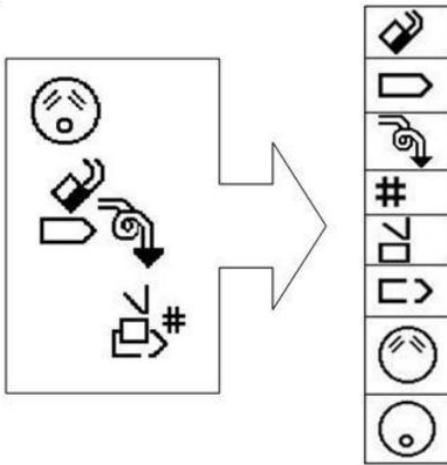


Figura 14: Exemplo de ordem de leitura/soletragem de um sinal. Os símbolos à direita estão ordenados pela ordem de leitura.
 Fonte: Moemedi, Connan (2010, p.2)

Durante a escrita pode ocorrer um deslocamento de elementos da quarta sílaba, como detalhamento da face e pontos de articulação, que ganham prioridade em relação à configuração de mãos (WANDERLEY, 2011). Este deslocamento reflete a relevância dos movimentos faciais na compreensão dos sinais.

Já a *sequência de símbolos do sinal* refere-se à ordenação de um sinal em um dicionário ou glossário de escrita de sinais. Esta sequência permite ao usuário encontrar os sinais nesse dicionário, baseando-se na configuração de mão ou outros detalhes (SUTTON, 2008). As categorias de símbolos do *SignWriting* podem ser vistas no quadro a seguir.

Quadro 2: categorias de símbolos do *SignWriting*
 Fonte: Denardi (2006, p. 29).

	Category 1: Hand
	Category 2: Movement
	Category 3: Face
	Category 4: Head
	Category 5: Upper-Body
	Category 6: Limb
	Category 7: Full-Body
	Category 8: Location
	Category 9: Dynamics
	Category 10: Punctuation

O *SignWriting* é um sistema de escrita fonética de grande valia para os aprendizes de línguas de sinais, entretanto o detalhamento de todos os fonemas no registro escrito pode limitar seu uso cotidiano, pois torna a escrita volumosa e complexa em função da quantidade e variedade de elementos gráficos (CORREIA; LIMA; LIMA, 2010).

Espera-se que o uso dos sistemas de escrita de sinais na legendagem possibilite a prática por parte do público surdo de sua língua na modalidade escrita.

A próxima sessão apresenta os ambientes em que o audiovisual pode ser acessado pelos públicos surdos.

2.2 AMBIENTES DE ACESSO

O acesso ao audiovisual digital pode ocorrer por diferentes meios, tais como IPTV, WebTV, Internet TV, TV Digital terrestre, TV Digital via satélite e P2P TV. Estas designações se referem ao meio de transporte de dados e pouco diferem quanto aos serviços que podem oferecer diante um cenário de convergência (SOARES et. al, 2010).

No contexto brasileiro, o censo de 2010 (IBGE, 2010) constatou que do universo de 57 milhões de domicílios amostrados, 54,5 milhões possuem aparelho de TV e 21,93 milhões possuem computador, sendo que destes, 17,6 possuem acesso à internet. Entretanto deve-se notar que a penetração da TV digital não é tão ampla, pois o censo não diferencia os tipos de televisores.

Em relação à oferta de audiovisual acessível aos surdos, no Reino Unido o objetivo da legislação é que os radiodifusores ofereçam, até 2013, cinco por cento dos programas com língua de sinais, seja pela tradução de programas de ouvintes ou pela produção de programas feitos por surdos. Além disso, 80% dos programas deverão ser legendados na língua oral. Nesse sentido, o site da BBC (BBC, 2011) possui uma área dedicada a língua de sinais que permite o acesso a programas sinalizados, conforme visto na figura 15.

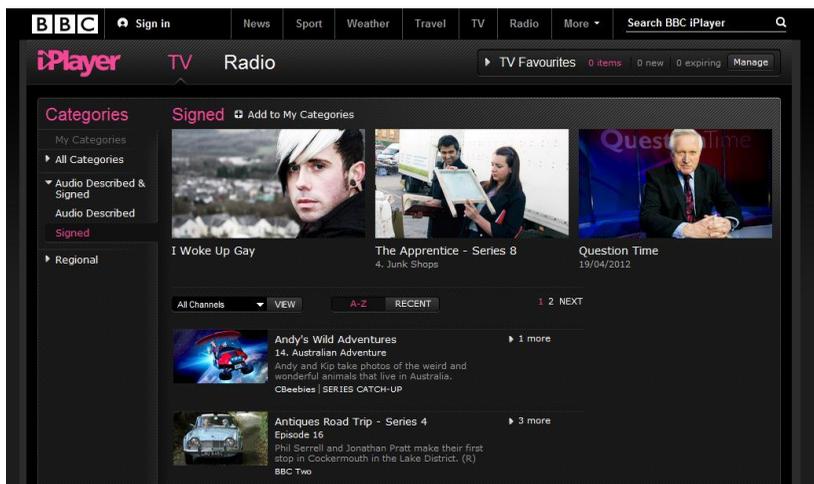


Figura 15: área de web site de emissora dedicada à oferta de conteúdo acessível.
Fonte: <http://www.bbc.co.uk/iplayer/tv/categories/signed>

A TV francesa oferece telejornais com tradução em língua de sinais, e também com legendas em língua oral. Exemplos são apresentados nas figuras 16 e 17.



Figura 16: TILS em noticiário francês.



Figura 17: TILS acompanhado de legenda textual.

A Figura 18 apresenta um exemplo de iniciativa de canal de entrega de audiovisual com acessibilidade aos surdos usuários de LS na web.



Figura 18: iniciativa de TV interativa na web “surd@ TV Brasil”

Fonte: www.stvbr.com.br

Ainda no contexto da web, a Figura 19 apresenta um exemplo de situação onde é possível a transcrição do áudio, por meio de reconhecimento de voz, e sua posterior tradução automatizada.



Figura 19: interface web para apresentação de vídeo com opção transcrição automática de áudio para geração de legendas em texto.

Fonte: YouTube

Para disponibilizar conteúdo acessível na web, Debevc *et al.* (2009) apresentam um módulo para apresentação de conteúdo de forma multimodal, englobando as modalidades visual, auditiva e textual. Na modalidade visual é apresentado um intérprete de língua de sinais, na modalidade auditiva é tocado um áudio com voz e a modalidade textual apresenta legendas. Um exemplo do módulo é apresentado na Figura 20.

The screenshot shows the FERI website interface. At the top, there is a blue header with the FERI logo and a small flag icon. Below the header is a navigation menu with links: Overview, Demo, Lab, Project, Authors, Publications, Contact, SLI Module, and E-learning portal. The main content area features a video player with a woman signing. The video content includes the following text:

1. Install the latest Adobe Flash Player: Adobe Flash Player 10.1+
 2. Enable JavaScript.
 3. Click a hyperlink/icon below.

LAYERED

Besedni zaklad
 Besedni zaklad so vse besede, ki jih nekdo pozna in jih uporablja.

Ustvarjaljen
 Človek je ustvarjaljen, kadar dela tako, da nastane nekaj novega. Ustvarjalni smo, kadar na novo risbo ali napišemo pesem.

Zanesljiv
 Človek je zanesljiv, kadar nimamo o njem nobenega dvoma, da bo na primer nekaj naredil.

Features:
 → Movable video
 → Activate/Deactivate video

Browser Compatibility:
 → Firefox 3.5+
 → Internet Explorer 7.0+
 → Chrome 5.0+
 → Opera 10.0+
 → Safari 4.0+

Platform:
 → Windows OS
 → Mac OS
 → GNU/Linux OS

Človek je zanesljiv, kadar nimamo o njem nobenega dvoma,

E-learning portal for deaf and hard of hearing

Figura 20: exemplo de aplicação de modulo de apresentação de LS
 Fonte: <http://slimodule.com/showcase.html> baseado em Debevc et al. (2009)

O módulo utiliza vídeos com fundos transparentes, de modo que o conteúdo continua visível durante sua exibição, sem alterar a estrutura da pagina web e permitindo o posicionamento do vídeo sobre qualquer lugar da tela. Esta abordagem deve ser utilizada com cautela pois o fundo do conteúdo pode prejudicar a visualização dos sinais. Na interação com o módulo, o usuário pode ativar os vídeos em LS sob demanda, sendo apresentados em sincronia com o áudio e as legendas.

No contexto nacional, a EBC por meio da TV Brasil, disponibiliza na web o Jornal Visual, um telejornal diário dedicado aos surdos usuários de Libras. A figura 21 apresenta o web site do Jornal Visual.

The image shows the TV Brasil website interface for 'Jornal Visual'. At the top left is the TV Brasil logo. A search bar with the text 'BUSCAR' is on the top right. Below the search bar are navigation links: 'Programação', 'Programas', 'Novidades', 'WebTV', 'Institucional', and 'Contato'. The main header features the 'Jornal Visual' logo in a large, stylized font. Below the header is a navigation bar with links: 'Início', 'Episódios', 'Vídeos', 'Fotos', and 'Sobre'. The main content area is divided into two columns. The left column contains a video player for 'Quarta, 25/04/2012 - 07h50' of 'Jornal Visual 25/04/2012'. The video shows a male news anchor and a female sign language interpreter. The video player has a progress bar at 0:52 / 9:55 and a volume icon. The right column has a 'Horário de exibição' section with the text 'Nacional Segunda a Quarta às 07h50'. Below that is a 'Vídeos' section with two video thumbnails: 'Jornal Visual 25/04/2012' and 'Jornal Visual 24/04/2012'.

Figura 21: telejornal dedicado aos surdos usuários de Libras
 Fonte: <http://tvbrasil.ebc.com.br/jornalvisual/>

Além da web, os conteúdos audiovisuais podem ser disponibilizados na plataforma de TV Digital Interativa (TVDI), cuja arquitetura é apresentada na Figura 22 a seguir. Na TVDI um programa interativo é formado por fluxos elementares de dados que contém a informação de áudio, vídeo ou dados de aplicativos, como os caracteres das legendas. Estes fluxos de dados elementares são multiplexados de modo a serem transmitidos em um meio comum e demultiplexados na recepção.

Após a demultiplexação cada fluxo elementar é remontado e apresentado por meio de decodificadores específicos, como no caso do áudio e do vídeo. Dados de aplicativos são encaminhados ao *middleware*, que constitui uma camada de software padronizada entre os receptores de TVDI.

No sistema de TV digital brasileiro, o *middleware* padrão é denominado Ginga, e apresenta recursos variados para o desenvolvimento de aplicativos interativos por meio de linguagens

declarativas (NCL), ou procedimentais (Java) (VEIGA; TAVARES, 2007).

Dados gerados no receptor, como respostas de enquetes ou pedidos de dados adicionais, podem ser enviados pelo canal de retorno, que constitui uma interface de comunicação com o emissor de conteúdo (KIM et al., 1998).

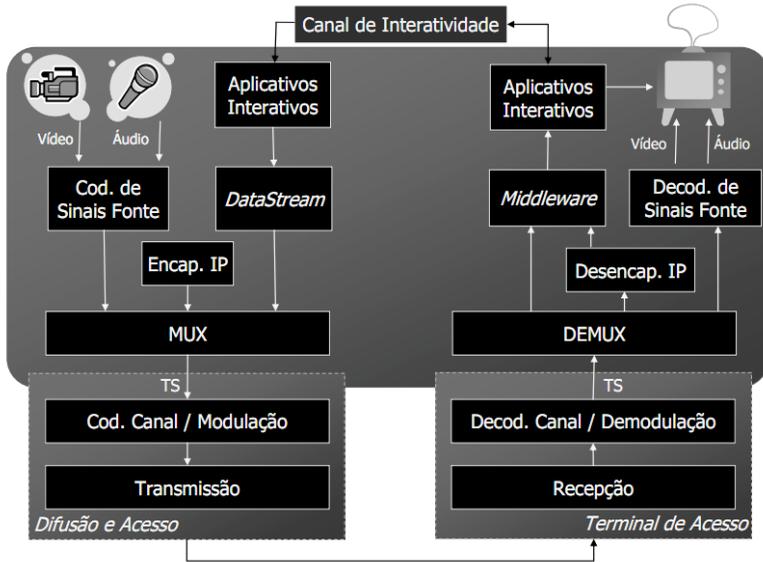


Figura 22: Sistema de TV digital
Fonte: Barbosa e Soares (2008)

Os serviços de TV digital podem ser classificados em dois tipos principais: os aplicativos vinculados ao conteúdo e os não vinculados (SOARES et al., 2010). Pode-se afirmar que as legendas para oferta de acessibilidade estão associadas ao conteúdo, de modo que estejam sincronizadas com as mensagens sonoras das cenas apresentadas.

A limitação da interatividade presente na TVDI, diante da interatividade presente nos computadores, se deve à imaturidade do setor e às restrições dos padrões técnicos (UGARTE et al., 2007).

O suporte a gráficos 3D nesta plataforma pode permitir a exploração de novas possibilidades na TVDI, entre elas o uso de avatares para apresentação de LS, entretanto projetar e implementar software baseado em gráficos 3D para hardware com capacidade limitada, como é o caso dos receptores de TVD, constitui uma tarefa

difícil, pois os cálculos para exibição de 3D requerem grandes quantidades de memória (ZHANG, 2003). Porém,

Felizmente, o hardware do set-top box vem sendo constantemente melhorado e se acredita que os set-top boxes são capazes de apresentar gráficos 3D em tempo real serão comuns no mercado, cedo ou tarde (ZHANG, 2003).

Atualmente o set-top Box tende a evoluir seu poder de processamento e se assimilar aos videogames, entretanto esta convergência ainda não aconteceu (ROBINSON, 2011).

O *middleware* do padrão europeu (MHP) pode oferecer serviços de acessibilidade, Martin et al. (2010) exploram sua aplicação na criação de um aplicativo de legendas para surdos e um Guia Eletrônico de Programação (EPG – *Electronic Program Guide*) para cegos. O EPG pode ser considerado um programa residente na memória do receptor de TVD, de tal forma que são pré-instalados nos receptores.

Outro tipo de aplicativos de TVDI, os não residentes, são enviados e instalados segundo pedidos do usuário final. A diferença prática entre estes dois tipos é apresentado por ABNT (2010):

As aplicações residentes podem ser implementadas usando funções não padronizadas, fornecidas pelo Sistema Operacional do dispositivo de Ginga [*middleware*], ou por uma implementação particular do Ginga. Os aplicativos residentes também podem incorporar funcionalidades providas pelas API padronizadas Ginga-J. Aplicativos transmitidos (*Xlets*) sempre devem utilizar API padronizadas fornecidas pelo Ginga- J (ABNT, 2010).

Para o empacotamento de conteúdos em Libras, Durand et al. (2005) descrevem um modelo de metadados, visto na figura 23, onde um programa (*Program*) é composto por itens (*Program ITEM*), o qual agrega um item de conteúdo principal (*Main Content Item*) e pode estar associado a itens de conteúdo adicionais (*Additional content Item*).

Estes itens possuem uma definição de tipo, que se enquadra em áudio, audiovisual, HTML, imagem ou vídeo, e têm um tipo semântico, que define sua função, que pode ser um comentário adicional ou então o conteúdo de tradução em língua de sinais (*signer*), por exemplo.

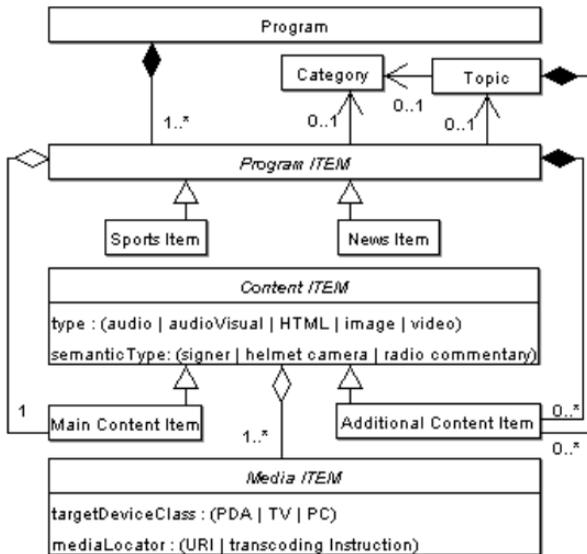


Figura 23: modelo de metadados para serviços de TV interativa
 Fonte: Durand et. Al (2005)

O modelo de metadados de Durand et al. (2005) tem o objetivo de descrever o conteúdo como contendo a tradução em língua de sinais, mas não predefine qual seu tipo entre vídeo, imagem, ou HTML. Os formatos de mídia para os artefatos de acesso ao audiovisual são apresentados no capítulo 5, que descreve o modelo de referência.

A próxima sessão explora questões associadas à tradução do audiovisual para surdos.

2.3 A TRADUÇÃO DO AUDIOVISUAL PARA OS SURDOS

A tradução é elemento chave para a oferta de acessibilidade aos surdos ao conteúdo audiovisual. Este capítulo descreve procedimentos de tradução para a produção de artefatos de apoio à acessibilidade de surdos ao audiovisual.

Nesse sentido, o audiovisual engloba os meios visuais e acústicos para a apresentação da informação (CHIARO, 2009), conforme descreve o quadro 3 abaixo. É possível classificar o uso de legendas de vídeo com intérprete de LS como informação de natureza visual e verbal, assim como as legendas textuais.

Quadro 3: A natureza polissemiótica dos produtos audiovisuais.

Fonte: Chiaro (2009, p.143)

	Visual	Acústico
Não verbal	Cenário, iluminação, roupas, expressões faciais, movimento corporal, etc.	Música, ruído de fundo, efeitos sonoros, etc.
Verbal	Sinais de rua, outdoors, placas, anúncios, etc.	Diálogos, letras de músicas, poemas, etc.

O tópico seguinte descreve as modalidades de tradução e os papéis envolvidos na oferta de acessibilidade considerando-se a variedade de meios envolvidos.

2.3.1 Tradução e Interpretação

Em um cenário em que o conteúdo audiovisual é produzido no contexto de uma língua oral, a acessibilidade deste conteúdo por surdos que utilizam uma língua gestual requer um processo de tradução.

A tradução e a interpretação são estratégias de apoio à acessibilidade ao audiovisual por meio da Libras. Na tarefa de tradução é possível um tempo para realização das adequações linguísticas entre a língua portuguesa e a LS, já na interpretação este processo ocorre de modo imediato, sendo que o emissor da mensagem é quem dita o ritmo de trabalho do Intérprete de língua de sinais (ALBRES, 2010). Neste trabalho a referência a este papel é definido como tradutor e intérprete de línguas de sinais (TILS).

O processo de tradução “consiste de escolhas, particular e é subjetivamente influenciada por aspectos ideológicos, linguísticos e culturais” (AUBERT, 1994 apud ALBRES, 2010, p. 137, tradução nossa), de modo que

o tradutor de Libras deve pertencer à cultura surda, ter boas raízes culturais e uma boa experiência na vida social em ambas as línguas; deve conhecer profundamente as várias nuances das duas culturas, encarando não só a estrutura linguística, mas também a vida cultural de uma sociedade como fatos semióticos (SEGALA, 2010, p. 8).

Para a tarefa de tradução, Segala (2010) descreve a aplicação de diferentes estratégias utilizando a frase em português:

A imaginação é mais importante que o conhecimento. (Albert Einstein) (SEGALA, 2010, p.47).

Tratam-se das estratégias de estrangeirização, domesticação e minorizante. Diante a estratégia de domesticação, a tradução da frase para Libras ficaria:

(ALBERT EISTEIN) FALOU: GRUPO AREA COISA GRUPO QUE É IMAGINAR IDEIAS COISA ESTE; OUTRO GRUPO AREA COISA CIENCIA ESTUDAR MENTE DESENVOLVER

COISA PESQUISAR COISA AREA; ESTE
(Imaginar) MAIS O-QUE IMPORTANTE DO-
QUE (ciência). (SEGALA, 2010, p.47)

A mesma frase traduzida para Libras pela estratégia de estrangeirização:

IMAGINAR MAIS IMPORTANTE DO-QUE
CONHECIMENTO. (ALBERT EISTEIN)
(SEGALA, 2010, p.50)

Ao se comparar as duas estratégias, a tradução domesticadora “facilita a compreensão, mas o surdo acaba perdendo o conhecimento da cultura e da língua da fonte original” (SEGALA, 2010, p.51), já na estrangeirização, “o surdo não consegue compreender, porque não tem conhecimento da cultura da fonte” (SEGALA, 2010, p.51).

Em complemento a estas duas estratégias, a estratégia de tradução ‘minorizante’ visa não reduzir a possibilidade de aprendizado do surdo nem o submeter a uma tradução que desconsidera sua cultura e língua. A frase de Einstein traduzida para Libras segundo esta estratégia é apresentada por Segala:

PESSOA ELE HOMEM NOME DELE A-L-B-E-
R-T E-I-S-T-E-I-N DIZER QUE GRUPO AREA
IMAGINAR I-M-A-G-I-N-A-Ç-Ã-O ESTE O
QUE IMAGINAR SONHAR IDEIA PENSA
ALVO FUTURO IDEIA PENSA ALVO COISA
IMAGINAR ESTE; OUTRO GRUPO COISA
CONHECER NOME C-O-N-H-E-CI-M-E-N-T-O
ESTE COISA DESENVOLVER CONHECER
PESQUISAR CIENCIA COISA ESTE;
COMPARAÇÃO (imaginação e conhecimento)
ESTE IMAGINAR MAIS IMPORTANTE DO-
QUE (conhecimento) (SEGALA, 2010, p.51)

Januário, Leite, Koga (2010, p.20) utilizam uma versão adaptada da pirâmide de *Vauquouis* para contextualizar a modalidade de tradução adotada, conforme a figura 24.

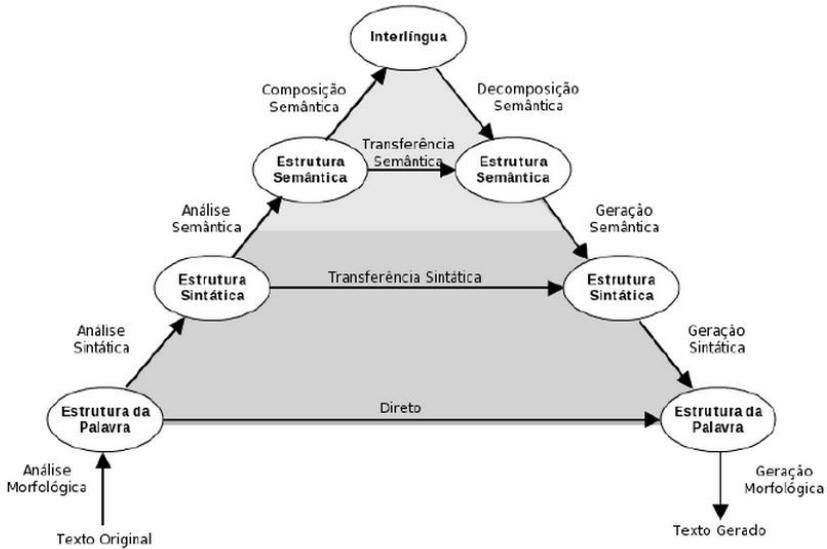


Figura 24: pirâmide de *Vauquois* adaptada por Januário, Leite, Koga (2010)
 Fonte: Januário, Leite, Koga (2010, p.20).

A pirâmide de *Vauquois*

apresenta níveis para a tradução, sendo que quanto mais alto na pirâmide, mais "profunda" é a análise feita na linguagem e portanto melhor seria a tradução realizada. No caso desta figura, a pirâmide original foi discretizada em quatro níveis. O primeiro nível, morfológico, representa uma tradução apenas baseada em palavras, ou seja, baseada apenas num dicionário. Esse tipo de tradução é pobre, pois desconsidera qualquer estrutura de frase ou relação entre as palavras. O segundo nível é o sintático, que realiza análise da sintaxe da língua, considerando então esses aspectos mencionados. O terceiro nível corresponde a um estudo das características semânticas do texto, levando em conta o significado das palavras, possibilitando a diminuição de ambiguidades presentes nos níveis abaixo e a identificação de expressões linguísticas. O topo da pirâmide seria a tradução perfeita, através de uma única língua intermediária na qual

as outras línguas pudessem ser descritas completamente (Januário, Leite, Koga (2010, p. 19).

Em relação à tradução, Jakobson (1995) identifica três tipos de tradução:

- a interlingüística, onde o texto de partida e chegada estão em línguas diferentes;
- a intralingüística, onde o texto de partida e chegada estão na mesma língua;
- e a intersemiótica, quando o texto de partida e de chegada estão representados em meios semióticos diferentes, do visual para o oral e vice-versa.

Estas modalidades de tradução estão presentes no processo de produção de legendas para surdos. Em termos gerais a tradução interlingüística tratará da questão da transposição de um texto da língua portuguesa para a Libras, já a questão intersemiótica é inerente à natureza sonora da língua oral que deve ser representada no meio visual e espacial das línguas de sinais, ou então sua tradução para notação escrita. Por sua vez a modalidade intralingüística ocorre, no contexto do audiovisual, nas reduções e expansões necessárias à manutenção da sincronia entre o conteúdo da tradução e o conteúdo original, ou então à adequação da quantidade de conteúdo à velocidade de leitura

A próxima sessão explora estas modalidades de tradução e seus relacionamentos em mais detalhes.

2.3.2 Modalidades de tradução aplicadas ao audiovisual

No contexto do audiovisual a tradução *intersemiótica* pode ser entendida como a codificação do conteúdo produzido em um meio sonoro ou visual para um código escrito ou vice-versa. É possível se falar em tradução intersemiótica quando ocorre a produção de legendas do tipo closed-captions, onde o áudio do vídeo em português deve ser representado em um meio escrito em sincronia com o conteúdo principal.

A tradução intersemiótica também ocorre quando o conteúdo original está na Língua Oral (LO), no caso o português, e passa a ser apresentado na Libras, numa representação gestual.

A tradução intersemiótica também é referenciada pelo termo ‘intermodal’, conforme afirma Segala (2010) quando se refere à tradução da Libras para o português.

A recodificação de uma mensagem originalmente produzida em Libras (língua gestual-visual) para o Português (língua oral-auditiva) enquadra-se no que vem sendo chamado de tradução intermodal (SEGALA, 2010, p. 27).

Abordagens de tradução intermodal apoiada por computadores são exemplificadas por Maurer, Stubenrauch e Camhy (2003):

Sistemas computacionais de apoio a tradução entre línguas orais e línguas gestuais - em ambas as direções - estão disponíveis em versões experimentais ou comerciais. Uma abordagem experimental de reconhecimento baseado em luvas de dados é apresentado por Kadous (1995), e uma abordagem baseada no reconhecimento por visão computacional é apresentado por Holden, (2001). A tradução da fala para os sinais é demonstrado no projeto da união europeia VISICAST e foi trazida ao mercado por meio de produtos como iCommunicatorTM (www.mycommunicator.com) e o SigningAvatarTM (www.vcom3d.com) (MAURER; STUBENRAUCH; CAMHY, 2003, p.335, tradução nossa).

Nascimento (2011) apresenta um estudo sobre a interpretação da Libras em telejornais onde identifica estratégias para produção de significados utilizados pelos intérpretes. De modo a oferecer acessibilidade ao texto jornalístico, é sugerida a aplicação do princípio da redundância, ou seja, oferecer meios multissensoriais para o acesso à informação, no caso, a apresentação de legenda escrita em português ao mesmo tempo que oferece a legenda com Libras.

Por sua vez, a tradução *interlingüística* consiste na conversão do conteúdo de uma língua para outra língua. Em relação ao audiovisual, a tradução interlingüística auxilia o espectador a entender o contexto de uma língua para outra (LU, 2010). Ela é aplicada na veiculação de conteúdos importados, como séries de TV ou documentários, produzidos em uma língua estrangeira, como o inglês ou espanhol, e traduzidos para a língua portuguesa. Nestes casos o conteúdo é traduzido de uma língua estrangeira para uma língua nativa.

A tradução interlingüística pode ocorrer sem que haja mudança do meio de comunicação, como na dublagem, em que a mensagem é codificada no meio sonoro.

Basicamente, quando se traduz um filme ou série de televisão, os tradutores tendem a retratar à sua audiência o espírito e essência do trabalho, ao invés de esforçarem-se em fornecer uma tradução literal e pura das palavras faladas. Já em programas de notícias e documentários, o tradutor prefere permanecer próximo ao diálogo original de modo a assegurar que sua audiência receba os fatos essenciais, de modo que preocupações estéticas assumem menor importância (PETTIT, 2004, p. 37, tradução nossa).

A tradução para legendagem em texto “consiste em passar um discurso oral para uma ou duas linhas escritas, ao mesmo tempo que se passa de uma língua para outra”, onde ocorre “a passagem de um código oral a outro escrito, para além da tradução interlingüística propriamente dita” (FERNANDES, 2003, p.39, tradução nossa).

No caso de tradução do português para a Libras, além de uma tradução intersemiótica, ocorre a tradução interlingüística. Cabe ressaltar que

O Tradutor intermodal e intersemiótico /interlinguístico deve conhecer as duas línguas e suas articulações em duas modalidades diferentes. O desconhecimento ou a falta de fluência pode comprometer as traduções e resultar na insatisfação dos leitores usuários da Língua Brasileira de Sinais (SEGALA, 2010, p.30).

Na tradução de audiovisual, é necessário o acesso prévio do intérprete aos materiais que serão apresentados de modo que este tenha uma referência para a interpretação. Além disso é comum o uso de recursos de expansão e compressão durante a interpretação de modo que a fala do intérprete seja simultânea com a do locutor (NASCIMENTO, 2011). Para a expansão ou compressão do conteúdo apresentado, a tradução *intra*linguística mantém a mesma língua e o mesmo meio de representação da mensagem na origem e no destino da tradução.

É possível analisar a tradução intralinguística nos meios de representação oral, gestual e escrito. O objetivo deste tipo de tradução no contexto da legendagem do audiovisual é condensar o conteúdo de modo a torná-lo mais acessível às audiências. A tradução intralinguística auxilia surdos ou aprendizes de uma segunda língua e ocorre dentro da mesma língua, e no caso dos surdos, consiste em uma estratégia interessante quando se simplifica a linguagem utilizada pelos ouvintes de modo a facilitar a leitura da mensagem.

Na tradução intralinguística, a edição pode ser entendida como “não dar Surdos tudo o que é dado aos ouvintes”, pois transcrever cada palavra, não serve às necessidades desse público específico. Deve-se remover a informação supérflua, introduzir notas explicativas, tornar explícito o implícito (NEVES, 2005).

Há também uma adaptação da linguagem diante a variação linguística decorrente dos espaços geográficos, dentro da própria Libras, conforme destaca Avelar (2010) a respeito da tradução em um cenário de ensino a distância.

deve haver respeito à variação linguística de cada ator-tradutor surdo, nas situações cotidianas de interação e também quando traduz livros especificamente para a região de onde é. Mas [...] é preciso levar em consideração que os atores-tradutores surdos são contratados para atenderem a um público bastante diversificado e têm aí a oportunidade de introduzir sinais novos na língua

para todos os alunos do Brasil. Por essa razão, os sinais devem estar padronizados (AVELAR, 2010, p.71).

O quadro a seguir sintetiza as diferentes modalidades de tradução onde se desenvolvem as estratégias de oferta de acesso ao audiovisual. O foco está na tradução do português para a Libras ou então na tradução intralinguística da Libras, de modo que as lacunas referentes à tradução da Libras para o português não são analisadas.

Quadro 4: modalidades de tradução para o acesso do surdo ao audiovisual e artefatos resultantes

Fonte: os autores

		Português		Libras	
		Oral	Escrito	Gestual	Escrito
Português	Oral	Intralinguística Dublagem oral	Intersemiótica e intralinguística, produzindo legendas textuais em português a partir da informação sonora	Intersemiótica e interlinguística, resultando em legenda com intérprete de LS	Intersemiótica e interlinguística, gerando texto em escrita de sinais
	Escrito		Intralinguística, permitindo a criação de legendas com diferentes níveis de condensação de conteúdo	Intersemiótica e interlinguística, resultando em vídeo com intérprete de LS	Interlinguístico, gerando texto em escrita de sinais
Libras	Gestual			Intralinguístico, refletindo diferenças de interpretação devido às diferenças regionais, resultando em vídeo com intérprete de LS	Intralinguística e intersemiótica, produzindo legendas textuais em sistema de escrita de sinais
	Escrito				Intralinguística que gera legendas multiníveis em escrita de sinais, podendo haver variação no sistema de escrita utilizado

Os artefatos resultantes das possibilidades de tradução descritas acima são o vídeo com intérprete, a legenda textual em português e a legenda em escrita de sinais.

2.4 CONCLUSÕES SOBRE O ACESSO DE SURDOS AO AUDIOVISUAL, AMBIENTES DE ACESSO E TRADUÇÃO

Nota-se grande diversidade de tipos quando analisamos a população dos surdos.

Em termos gerais existem os sujeitos que ficaram surdos antes do desenvolvimento de uma linguagem e os surdos pós-linguísticos, que ficaram surdos após o desenvolvimento de uma linguagem para comunicação. Surdos pós-linguísticos tendem a ter boa afinidade com a língua oral e por consequência o português escrito, já os surdos pré-linguísticos tendem a se familiarizarem mais com as línguas de sinais. Entretanto estas associações não podem ser levadas como regra.

Muitos fatores culturais influenciam a constituição do surdo e o desenvolvimento de suas estratégias para comunicação, tais como o ambiente familiar e escolar. Diante estes fatores culturais, a abordagem da educação bilíngue visa possibilitar a comunicação dos surdos por meio de português, escrito ou oral e por meio de língua de sinais, respeitando a língua materna do sujeito e sua cultura.

A educação bilíngue com o respeito à cultura das línguas de sinais mostra-se mais adequadas à realidade do surdo, sobretudo ao enxergarmos o surdo sob um foco que não está na ausência de determinada habilidade sensorial, mas na complementaridade que existe entre o mundo surdo e o mundo ouvinte.

Por sua vez a escrita das línguas de sinais forma uma das bases para o registro e comunicação de uma identidade cultural surda, e demonstra relevância que será considerada na oferta de audiovisual acessível. Ressalta-se que o uso efetivo da escrita de sinais depende da inclusão desta modalidade de escrita na educação dos sujeitos surdos.

O acesso de surdos ao audiovisual depende de artefatos que permitam o desenvolvimento de suas estratégias de acesso. Estas estratégias são tão variadas quanto os fatores que influenciam a formação do surdo. Diante a caracterização do público surdo, percebe-se que estes são melhor inseridos na sociedade ouvinte quando esta inclusão ocorre sob a perspectiva de uma cultura bilíngue, de tal modo que a oferta de acessibilidade ocorre por meio dos artefatos de legendas em português e de línguas de sinais, tanto em vídeo quanto na forma escrita.

A oferta de audiovisual acessível baseia-se na tradução, de modo que as modalidades intersemiótica, interlinguística e intralinguística estão presentes na promoção de acessibilidade do audiovisual. O

mapeamento entre estas modalidades de tradução permite identificar as possibilidades de artefatos para o audiovisual acessível.

Por sua vez, a caracterização dos ambientes de acesso possibilita identificar as restrições ou possibilidades de cada plataforma. A web e a TVDI são meios de entrega do audiovisual, com os quais os recursos de acessibilidade devem ser compatíveis de modo a possibilitar sua entrega e utilização por usuários surdos.

Dois pontos podem ser enumerados em relação aos ambientes de acesso: a capacidade de processamento do hardware e o formato com que o conteúdo é entregue.

Quanto maior a capacidade de processamento do receptor do usuário, mais compacta poderá ser o formato de dados utilizado para entrega do conteúdo, reduzindo a demanda sobre o fluxo de dados transmitido.

No contexto da TV digital, que possui uma plataforma mais restrita, a entrega de legendas deve seguir padrões estabelecidos ou então basear-se em implementação de aplicativos residentes no aparelho receptor.

Já a web oferece maior flexibilidade quanto aos recursos que podem ser utilizados, tanto para a oferta de legendas escritas quanto em vídeo.

O próximo capítulo caracteriza os artefatos de apoio à acessibilidade dos surdos e apresenta os processos associados a sua produção.

3. ARTEFATOS DE APOIO AO ACESSO DOS SURDOS AO AUDIOVISUAL

Este capítulo descreve e analisa as características dos artefatos de apoio à acessibilidade de surdos e os processos e técnicas associadas a sua produção.

Existem, essencialmente, duas ferramentas que ajudam as pessoas com surdez a receberem os conteúdos veiculados na televisão. Uma é o *close-captions* e a outra é a janela com o intérprete de Libras (ALBRES, 2010, p.131).

Em uma análise mais ampla, os artefatos que podem apoiar o acesso dos surdos ao audiovisual são os resultantes dos possíveis processos de tradução descritos no capítulo anterior: a legenda em português escrito, em vídeo com intérprete e a possibilidade de uso da escrita de sinais.

As sessões seguintes apresentam os processos e artefatos voltados à acessibilidade de surdos que preferem o português escrito e os voltados aos surdos que utilizam as línguas de sinais.

3.1 LEGENDAS EM TEXTO NA LINGUA ORAL

A legendagem textual em língua oral (LO), no caso brasileiro em português, pode ser classificada em dois tipos básicos: a aberta e a fechada. As legendas abertas são parte integral do audiovisual e não podem ser removidas pelo espectador, já as legendas fechadas são apresentadas por meio de um decodificador e gerador de caracteres específico. As legendas fechadas são menos intrusivas do que as abertas, pois não obstruem a visualização por audiências ouvintes, sendo possível oferecer informações adicionais aos surdos (FERNANDES, 2003).

Deste modo, o uso de legendas fechadas, ou opcionais, é mais adequado e serão o tipo de legenda utilizado no modelo.

Cabe ressaltar que a transcrição da fala em legendas fechadas não atende plenamente a demanda de acessibilidade, como no caso do *closed-captions*, pois os

Telespectadores surdos se beneficiam de legendas que são sintática e semanticamente estruturados de modo a facilitar a compreensão. Sentenças longas e complexas irão certamente demandar mais da memória de trabalho de curto-termo. Estruturas simples e diretas, com pausas frasais adequadas (por exemplo, não separando em linhas diferentes os artigos dos substantivos) facilitam a compreensão e tornam a leitura das legendas muito mais efetiva (NEVES, 2005, p.149).

O texto é visto como uma forma nem sempre mais adequada de apresentar informação aos surdos nativos em línguas de sinais (FELS et al., 2005), sendo que “a legendagem é frequentemente considerada como não emotiva e inadequada para transmitir a mensagem plenamente” (DAVIES et al., 2007, p.1135).

Além disso, o recurso de *closed-captions*, utilizado para transmitir legendas na TV, apresenta problemas em relação ao posicionamento e rolagem do texto, atrapalhando o usuário na sua leitura. A figura 25 apresenta um exemplo de falhas na legenda em português, tanto na decodificação de caracteres quanto na sua digitação.

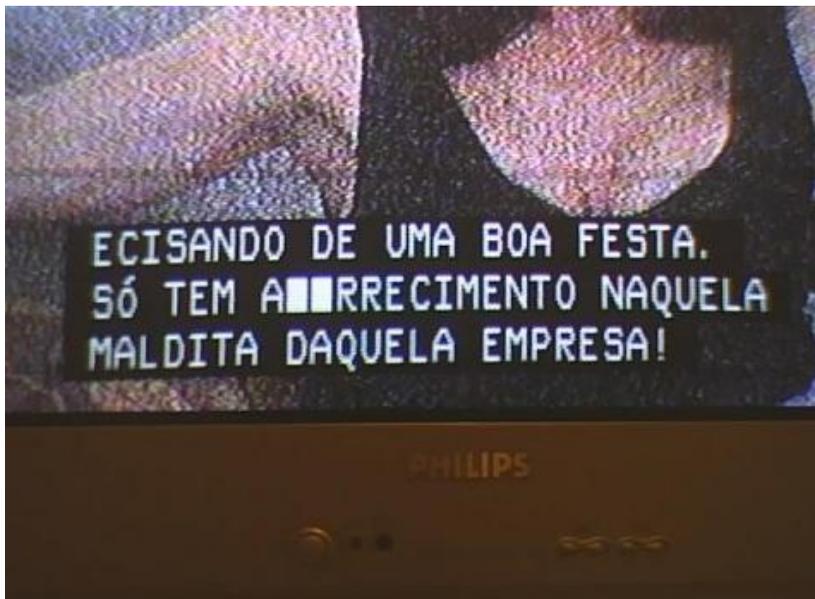


Figura 25: exemplo de problema com legendagem em *closed-captions* analógico

Fonte: <http://tadificil.wordpress.com/category/acessibilidade/>

Reichert (2006), ao analisar a mídia televisiva do ponto de vista de alguns surdos, afirma que

Para nós surdos, as legendas [em texto] são algo estranho, representam um mundo de vozes que não são perfeitamente audíveis. Os sons que as legendas significam não fazem parte natural de nossas vivências, não temos uma memória auditiva onde possam, confortavelmente, se instalar, são sempre arduamente armazenados por meio de uma série de estratégias de articulações de informações. A isto se soma a velocidade com que devem ser mostradas e substituídas (REICHERT, 2006, p. 66).

Deste modo,

É fundamental ter em mente que pessoas com surdez, e muito particularmente aquelas que ficaram surdas antes de adquirirem uma língua oral

(surdez pré-lingual), ou que não dominam a língua gestual, revelam lacunas linguísticas³ que se refletem numa leitura mais lenta e numa maior dificuldade na interpretação do texto escrito. Uma vez que a legendagem apresenta texto fragmentado, condicionado pela velocidade e ritmo do texto audiovisual, este exige um grande esforço de leitura seletiva e de memória. A pessoa com surdez beneficiará de qualquer estratégia que a ajude a desenvolver a sua capacidade de leitura, a ativar a sua audição residual e/ou a ultrapassar ruídos inerentes à sua condição (NEVES, 2007, p. 10).

Assim, as pessoas ensurdecidas ou que ainda tem uma audição residual, conseguem relacionar as legendas à fala, possuindo assim maior possibilidade de compreensão do conteúdo legendado. Porém no caso de pessoas com surdez pré-lingual, especificamente os surdos que tem a língua de sinais como língua materna, consideram difícil interpretar mensagens que estão representadas em um meio que não podem relacionar diretamente a seus sentidos (NEVES, 2005).

Este fato, entretanto não elimina os benefícios do acesso à legendas escritas na língua oral para surdos pré linguísticos, conforme afirma Reichert (2006):

O uso da legenda nos programas televisivos tem gerado outras necessidades para nós surdos. Se antes reivindicávamos o acesso e o desenvolvimento surdo em Língua de Sinais, hoje já tendo garantido o primeiro caso, começam a surgir outras necessidades, por exemplo, do aprendizado do Português escrito. Precisamos saber o significado das palavras para compreendermos o que está escrito nas legendas, cada vez mais somos chamados a conhecer mais palavras e, cada vez mais, estamos mais exigentes por qualidade de ensino de Português para surdos. Diante da leitura e das palavras desconhecidas notamos muitas estratégias de tradução e compreensão dos significados possíveis para as mesmas, às vezes, anotamos as palavras

³ A lacuna é em relação à língua oral e não necessariamente uma lacuna linguística.

desconhecidas em um caderno para buscarmos na internet ou no dicionário um sentido para ela, outras vezes perguntamos para algum ouvinte que está próximo ou anotamos e esperamos que a mesma palavra apareça em outros momentos para irmos montando significados para ela. (REICHERT, 2006, p.51)

De modo a classificar as legendas textuais, Teixidor (2008) descreve parâmetros para uma taxonomia de legendagem, os quais podem ser divididos em linguísticos, pragmáticos e técnicos. Os parâmetros linguísticos referem-se à tradução ocorrida na legendagem, podendo ser em relação à língua em si (interlinguístico ou intralinguístico) ou em relação à densidade do conteúdo que foi trabalhada do processo de tradução: se reduzido ou se foi mantido o conteúdo integral.

Os parâmetros pragmáticos identificam aspectos da legendagem referentes a seu público alvo (se surdos ou ouvintes, por exemplo), à intenção da legendagem (se documental, quando voltada à comunicação, ou instrumental, como no caso das legendas para a aprendizagem de outra língua) (WILLIAMS; THORNE, 2000; LEWIS; JACKSON, 2001); ao tempo de sua produção (se anterior ou simultânea, como no caso de programas ao vivo) e à autoria (se feita por profissionais, por amadores ou de forma automática).

Os parâmetros técnicos estão associados à opcionalidade da legenda (se a mesma pode ser desativada, como no caso das legendas fechadas), a sua forma de difusão (projetada, como no cinema, ou emitida, como na televisão), às cores (se poli ou monocromática), a sua incorporação (se a legenda aparece de forma estática ou dinâmica, com rolagem) e a seu arquivamento (se a legenda pode ser armazenada de forma independente do conteúdo).

Um exemplo de classificação da legenda em português para o audiovisual televisivo acessível aos surdos, segundo os parâmetros de Teixidor (2008), é apresentado no quadro 5:

Quadro 5: Classificação da legendagem em português de acordo com os parâmetros de Teixidor (2008).

	Parâmetros	
--	------------	--

Linguísticos	Tipo de tradução	Intralinguística e intersemiótica, do português oral para o português escrito
	Densidade	Redução ou integral
Pragmáticos	Destinatário	Surdos
	Intenção	Documental ou instrumental
	Tempo	Anterior ou simultânea
	Autoria	Profissional, ou amadora, ou automática
Técnicos	Opcionalidade	Opcional (legenda fechada)
	Difusão	Emitida
	Cor	Policromática
	Incorporação	Estática
	Arquivamento	Em separado

O controle do posicionamento e do tamanho das legendas é uma funcionalidade desejada pelos usuários em geral (VY; FELS, 2009, 2010). Também o uso dos nomes dos personagens na legenda auxilia no entendimento do conteúdo pelos surdos. Entretanto o posicionamento da legenda de acordo com o personagem distrai o usuário (VY; FELS, 2010).

O tópico seguinte apresenta as alternativas de processos para produção de legendas em português.

3.1.1 Produção de Legendas Textuais em Língua Oral

A produção de legendas em texto na LO é a abordagem mais representativa na oferta de acessibilidade ao audiovisual, entretanto

as práticas e processos associados à legendagem não mudaram nos últimos 30 anos. O propósito do uso de letras em caixa alta está relacionado à inabilidade da tecnologia dos anos 70 em mostrar caracteres descendentes, como o ‘g’ e o ‘y’. Entretanto a tecnologia televisiva avançou consideravelmente desde então [...], e apesar disso as práticas de legendagem permanecem essencialmente as mesmas, mesmo que haja ampla evidência de que a apresentação de letras com caixa variável facilita a leitura (UDO, FELS, 2010, p. 216, tradução nossa).

Na intenção de melhoria das práticas associadas à legendagem, Neves (2005) analisa a legendagem em português escrito na apresentação do audiovisual para surdos e pessoas com dificuldades auditivas. A partir de sua análise são definidas recomendações para a produção de legendas textuais, um desenvolvimento sugerido é a legendagem multinível ou multitrilha, a qual permite o espectador determinar as características das legendas que quer receber.

Isto significa a possibilidade de escolha entre um número de tipos graduais de legendas, partindo de uma legendagem textual direta até legendas editadas e adaptadas às necessidades de diferentes leitores. Adicionalmente, a possibilidade de escolha permitirá aos espectadores definirem qual tipo de conteúdo desejam incluir em suas legendas: informações extras sobre efeitos sonoros e música, cores ou posicionamento, entre outros. O tamanho das legendas também deve ser modificável, atendendo as necessidades de pessoas que, além da imparidade auditiva, pode também possuir baixa visão, uma condição comum em uma sociedade que está envelhecendo (NEVES, 2005, p.316, tradução nossa).

As mensagens textuais apresentadas possibilitam aos surdos uma melhor recepção de programas legendados, entretanto elas devem ser adequadas a estes sujeitos (POLLARD et al., 2009) e serem publicadas numa plataforma amigável (OKURA; HIROSE, 2008). Nesse sentido, a condensação e a edição das legendas são fundamentais para tornar mais

eficazes a leitura e compreensão do conteúdo por espectadores surdos (ARAUJO, 2008).

as legendas para surdos devem condensar aproximadamente 30% daquilo que é falado, ou seja, para cada 1s de fala, devemos traduzir o que foi dito com 14 caracteres (ARAUJO, 2008, p.74).

Nas legendas, vocabulários mais avançados devem ser utilizados apenas quando possuem um objetivo útil e deve ser oferecido tempo para a assimilação de seu significado pelo usuário (NEVES, 2005).

No caso da legendagem intralinguística, onde a transcrição literal da fala é frequentemente procurado, a simplificação do texto é particularmente difícil. Na legendagem interlinguística, as mudanças funcionais são menos expostas e, portanto, pode ser mais fácil adaptar o texto às necessidades dos destinatários surdos (NEVES, 2005, p.150).

Na intenção de permitir a automação da simplificação de frases do português, Caseli et al. (2009) e Candido et al. (2009) apresentam uma proposta de um sistema que visa oferecer acessibilidade de textos.

A redução seletiva da quantidade de palavras por minuto na legenda melhora a compreensão por partes dos leitores menos proficientes. Pessoas com dificuldade de audição e boa leitura se saem melhor a uma taxa de 180 palavras por minuto, enquanto que leitores surdos foram melhores com uma taxa menor, de 130 palavras por minuto (BURNHAM et al., 2008).

duas opções são deixadas como recomendação para uso em legendas: (a) selecionar taxas de legenda e métodos de redução de texto que sejam adequados à maioria dos espectadores na maioria das situações ou (b) oferecer a entrega de acordo com opções individuais” (BURNHAM et al., 2008, p. 402, tradução nossa)

A legendagem tradicional não apresenta informações importantes sobre aspectos emocionais do diálogo, música de fundo e identificação de personagem. Para apresentar estes aspectos na legenda, Fels et al.

(2005) constataram que sujeitos com dificuldade de audição preferem ocasionalmente o uso de gráficos, cores e ícones para apresentação da informação sonora, entretanto usuários surdos não aceitaram estes recursos como uma alternativa à legendagem tradicional.

Estas diferenças e preferências entre espectadores surdos e com audição prejudicada podem ser gerenciadas pela definição de preferências ou entrega de múltiplas trilhas de legenda de modo a permitir a escolha do espectador entre legendas com recursos gráficos ou legenda apenas com texto (FELS, et al., 2005).

Para a produção de legendas de texto oral, Neves (2005) sugere as seguintes etapas:

1. Checagem do script ou transcrição do diálogo.
2. Tradução e adaptação (transadaptação)
3. Sincronização (cueing)

Detalhadamente, o processo de produção de legendas para televisão, segundo Tveit (2004), envolve as etapas de:

1. Decodificação, que pode ser bastante difícil caso não se tenha a lista do diálogo e a qualidade do som seja ruim;
2. Tradução, que envolve encontrar os equivalentes na língua-alvo que são apropriadas em termos de correção e adequação;
3. Digitação do texto;
4. Condensar, deixando de lado as palavras menos importantes do texto de modo a facilitar a leitura;
5. Sincronização do texto com o tempo do vídeo;
6. Edição, que permite adicionar os últimos retoques ao texto, alterando a sua formulação, tempo e design gráfico.

As legendas também podem ser geradas a partir do reconhecimento automático de fala. Um sistema desenvolvido para tal finalidade utiliza o software IBM *ViaVoice* em conjunto com um módulo auxiliar para correção de erros de reconhecimento, e foi utilizado em três companhias de legendagem no Reino Unido (LAMBOURNE et al., 2004; BAIN; BASON; WALD, 2002).

O reconhecimento de texto a partir de áudio apresenta uma taxa de acerto de 80%, sendo que a cada hora de conteúdo traduzido automaticamente, são necessárias 3 horas para sua correção. De modo a

diminui este tempo, Wald (2006) apresenta um sistema de apoio a correção de legendas. Outro fator a respeito do uso de reconhecimento automático do áudio é o atraso no tempo de tradução, devido ao tempo necessário ao processamento, fato que pode ocasionar problemas na compreensão do conteúdo. De modo a tratar este problema, pode-se dispor do uso de um *Personal Video Recorder* (PVR) que permite gravar o conteúdo para sincroniza-lo com o texto das legendas (Rander; Looms, 2010).

Para as legendas de português em texto, a abordagem de legendagem multinível permite que o próprio usuário escolha os artefatos mais adequados à sua estratégia de leitura.

A oferta de legendas em português escrito permite o entendimento do conteúdo do audiovisual, e fomenta a cultura bilíngue entre os surdos, porém

as legendas automatizadas em português, em substituição aos textos sonoros produzidos pelos meios de comunicação via closed-captions ou mesmo legendas de melhor qualidade, produzidas para vídeos ou para filmes, não têm atendido de forma satisfatória a grande maioria da comunidade surda brasileira, que usa a Libras (Língua Brasileira de Sinais) como sua primeira língua (LIRA, 2009, sem paginação).

No contexto nacional “a maioria da população de surdos brasileiros tem dificuldades em ler e escrever o português” (MOTTA, 2005, p.7), e nem sempre o usuário prefere a língua oral, tendo preferência pelas legendas em línguas de sinais.

Deste modo a próxima sessão descreve o desenvolvimento de legendas com intérprete de LS.

3.2 LEGENDAS COM INTÉRPRETE DE LÍNGUAS DE SINAIS

A comunidade surda tem se manifestado pela preferência ao uso da língua de sinais, visto que a presença majoritária de legendas textuais não torna acessível o audiovisual aos surdos não alfabetizados na língua oral (KURZ; MIKULASEK, 2004). Grande parte da população surda não conhece a escrita da língua oral ou possui dificuldade em sua aprendizagem (HUENERFAUTH, 2005). Cabe ressaltar que a aprendizagem da língua oral escrita para alguns surdos constitui um desafio diante a ausência de referenciais de memória auditiva em alguns casos. Entretanto esta aprendizagem potencializa sua participação na sociedade do conhecimento que possui uma natureza predominantemente ouvinte.

Há diferença na forma de organização do pensamento quando se observa a língua oral e a LS, e esta tradução entre os modos de pensamento é um dos desafios que os tradutores/intérpretes enfrentam quando precisam produzir conteúdo acessível por meio de legendas. As legendas devem atender requisitos de sincronia temporal para que o conteúdo seja compreensível e ao mesmo tempo não seja suprimido.

Os intérpretes de língua de sinais se interpõem entre línguas e culturas, como mediadores que assumem o lugar de tensão das diferentes vozes, a dos surdos, dos ouvintes e de sua própria voz. Um dos principais dilemas do intérprete de língua de sinais, ao intermediar uma relação no ato tradutório, gira em torno do conflito de representação fidedigna do conjunto dessas vozes, e conseqüentemente sua responsabilidade com a tradução (MASUTTI, 2007, p.144).

Os Tradutores e Intérpretes de Língua de Sinais nos programas televisivos possibilitam aos surdos viverem emoções que não podem ser experienciadas quando apenas a legenda em português está disponível (REICHERT, 2006).

O intérprete de Libras tem a especificidade de traduzir a Língua Brasileira de Sinais para a Língua Portuguesa e vice-versa, tornando-se o canal de comunicação entre falantes de uma e de outra língua. Este intérprete pode atuar em diferentes contextos comunicacionais como, por

exemplo, em palestras, programas de televisão, salas de aula, cultos religiosos etc. (CORREIA, MACEDO, LIMA, 2008, p.2)

Exemplos de janelas com intérpretes de línguas de sinais são apresentados nas figuras 26 a 29 a seguir.



Figura 26: exemplos de apresentação da língua de sinais em conteúdo audiovisual na TVDI.

Fonte: Certic (2007, p.6)

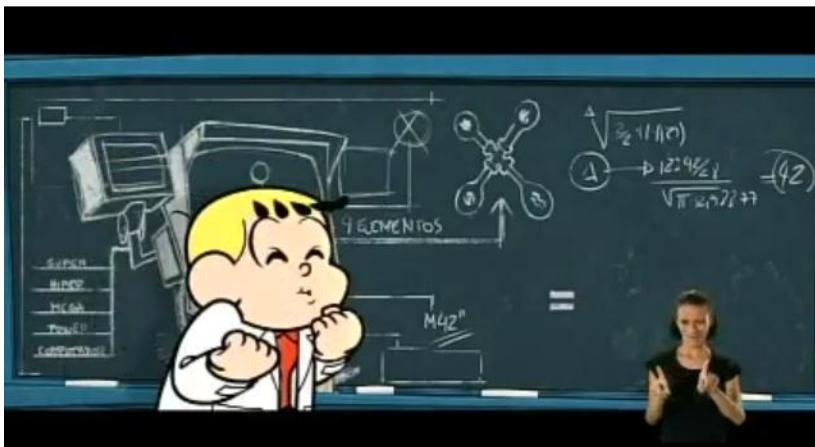


Figura 27: Exemplo DVD da Monica "Uma Aventura no Tempo"

Fonte: <http://www.youtube.com/watch?v=zFCV1Vkl71g>



Figura 28: Exemplo vídeo do YouTube

Fonte: <http://www.youtube.com/watch?v=2TbuksmNsJY>



Figura 29: Exemplo de legenda com Libras em conteúdo ao vivo pela web e TV

Fonte: NBR em 07/04/2011

Podem ser identificados essencialmente dois tipos de recorte de janela para inserção do TILS: o recorte em *Chroma Key*, conforme figura 27, e o recorte em janela, conforme figura 29 (NASCIMENTO, 2011). O recorte em janela oferece melhor contraste entre o intérprete e o fundo, oferecendo maior legibilidade dos sinais, já o recorte em

Chroma Key permite melhor integração do intérprete com a composição do audiovisual, porém elementos do fundo da cena podem eventualmente prejudicar a legibilidade dos sinais.

Fajardo, Vigo e Salmeron (2009) analisam tecnologias de apoio às línguas de sinais e propõem uma taxonomia para classificação de suas dimensões técnicas e funcionais. A taxonomia de Fajardo, Vigo e Salmeron (2009) utiliza as seguintes dimensões para classificação das tecnologias para produção de conteúdo com LS:

- Modalidade de comunicação, se escrita ou gestual;
- Gravação, se pré-gravada, ao vivo ou orientada ao usuário;
- Tarefa, se para busca ou navegação, comunicação ou aprendizagem;
- Teste em usuários, indicando sua aplicação prática e aceitação;
- Aplicabilidade em outros dispositivos, como celulares, televisores ou computadores, dispositivos com diferentes dimensões de tela e capacidade computacional.

Caso a tecnologia esteja na modalidade de sinalização gestual, outras dimensões podem ser usadas em sua caracterização:

- Espaço de sinalização, se inclui todos os aspectos necessários à comunicação por LS, se os sinais são apresentados em corpo inteiro ou apenas pela parte superior do corpo, se apresenta expressões faciais, posicionamento, movimentação, orientação e forma da mão, postura corporal e uso do espaço em torno do sinalizante;
- Tipo de transmissão, se simultânea ou em separado do conteúdo principal, indicando se os sinais são apresentados conjuntamente ao conteúdo hipermídia ou se apresentado em popups, por exemplo;
- Sinalizante, se humano ou avatar;
- Outros, como o apoio de ferramenta de autoria ou se é específica a uma LS, por exemplo Libras ou ASL.

Quando a comunicação for gestual e as técnicas utilizadas são baseadas em avatares, estas possuem quatro dimensões adicionais:

- Quanto ao tipo de avatar utilizado, podendo ser avatares humanoides realísticos (HUENERFAUTH, 2008) ou um Cartoon (YANAGIMOTO, 2004);
- Dimensionalidade, que pode ser 2D ou 3D, havendo preferência pela 3D por permitir melhor compreensão dos sinais;
- Quanto à localização da renderização, no contexto da web, esta pode ser do lado servidor ou do lado cliente, gerando vídeos pré-gravados ou sob demanda;
- Tecnologia gráfica, como VRML, X3D, bibliotecas gráficas ou representações abstratas.

O quadro 6 apresenta os parâmetros e as possíveis opções para classificação da tecnologia para produção de conteúdo em LS.

Quadro 6: Classificação de tecnologias para produção de conteúdo em LS

Fonte: Fajardo, Vigo e Salmeron (2009)

	Parâmetros	Opções
Gerais	Modalidade	Notação escrita ou comunicação sinalizada
	Tarefa	Busca, navegação, comunicação ou aprendizagem
	Geração (gravação)	Pré-gravada, ao vivo ou orientada ao usuário
	Aplicação em outros dispositivos	TV digital, celular, computadores pessoais
	Teste por usuários	Nenhum, satisfatório, requer revisão ou insatisfatório
Comunicação sinalizada	Sinalizante	Humano ou avatar
	Espaço de sinalização	Face, mãos, braços, tronco, espaço vizinho, parte superior do corpo
	Transmissão da LS	Simultânea ou em separado
	Outros	Apoio de ferramenta de autoria Especifica a uma LS Apoio de ferramenta de revisão
Uso de avatar	Tipo de avatar	Realista, Humanoide ou Cartoon
	Dimensões	Tridimensional ou Bidimensional
	Localização da renderização	Server-side ou client-side
	Tecnologia gráfica	DirectX, OpenGL, VRML, X3D, h-Anim

Das duas abordagens principais para a apresentação de LS na web: a partir de vídeo de humanos sinalizantes ou a partir de avatares, os vídeos de sinalizantes humanos são mais ‘ecológicos’, mas requerem um grupo de especialistas de alto custo para a realização da tradução, edição de vídeo e upload de conteúdo. Já a tecnologia baseada em avatares diminui os requisitos em relação aos recursos humanos e sobre o canal de transmissão, entretanto sua apresentação está geralmente associada ao baixo realismo e com a tradução automática, fato que reduz sua confiabilidade (FAJARDO; VIGO; SALMERON, 2009).

Em relação à modalidade de comunicação, esta pode ser por notação escrita dos gestos, como a HamNoSys e Stokoe, por comunicação sinalizada entre sinalizantes humanos, por meio de vídeo pré-gravado, e por comunicação sinalizada por avatares sinalizantes: que utiliza os sistemas de notação para a síntese dos sinais no navegador web do usuário. Esta abordagem tem o realismo, expressividade e legibilidade reduzidos, sendo que também a notação escrita dos sinais não é muito popularizada.

As ferramentas de autoria podem oferecer meios para produção de conteúdo, utilizando uma linguagem de marcação declarativa ou vídeos pré-gravados.

A tradução de uma língua oral para uma Língua de Sinais pode ser feita a partir de um sistema de escrita ou de vídeo (Segala, 2010).

Deste modo, os próximos tópicos analisam a produção de legenda em LS que utilizam vídeo e as que usam os sistemas de escrita de sinais. Tópicos subsequentes descrevem os sistemas de tradução automática e o uso de avatares, assim como os glossários/dicionários que apoiam o acesso aos artefatos produzidos.

3.2.1 Produção de Legendas com Intérprete de Libras com Vídeo

A produção de conteúdo acessível em Libras é intensiva em recursos, pois requer equipamentos de gravação de vídeo, edição e renderização, assim como recursos humanos, como intérpretes e tradutores, sendo que estes são as figuras chaves no processo de produção do audiovisual acessível.

Em relação ao desenvolvimento de legendas com LS no audiovisual, Albres (2010) identifica diferentes denominações para esse tipo de legenda.

As empresas que trabalham com produção de vídeos, utilizam a imagem em vídeo sobreposta à campanha publicitária em televisão e a denominam de ‘legenda em língua de sinais’, ‘legenda quadrada’ ou ‘legenda de intérprete de Libras’ (ALBRES, 2010, p.134).

Na produção de legendas em LS o vídeo permite manter o aspecto do intérprete o mais próximo ao real. Três etapas podem ser identificadas para a produção de material com vídeo em LS (JISC, 2010):

1. Pré-produção, com a preparação do vídeo que deve ser interpretado;
2. Produção, onde ocorre a gravação do intérprete contra uma tela verde, de modo que este possa ser inserido no vídeo por meio de técnica de *chroma-key*;
3. Pós-produção, quando se adiciona a imagem do intérprete no material de vídeo original e estes são codificados para entrega.

Para a gravação de vídeos, recomenda-se que o conteúdo seja filmado em unidades pequenas, denominadas *videolets*, com no máximo 5 minutos de duração, que podem ser concatenadas para formar vídeos mais longos. Esta prática visa facilitar a revisão do material produzido e reduzir a quantidade de vídeo que precisa ser refilmada no caso de modificações. (FELS et al., 2009).

A gravação das *videolets* é feita a partir de um script e pode utilizar os métodos *Slideshow*, *Serial Relay* e o *Parallel Relay* (FELS et al., 2009).

No método *Slideshow*, uma ferramenta de apresentação de slides é usada para apresentar o texto ao sinalizante que esta sendo filmado, sendo que para cada *videolet* um ou mais slides são utilizados.

No método *Serial Relay*, o sinalizante realiza os sinais para a câmera, com ou sem notas e o arquivo de vídeo resultante é editado posteriormente.

O método *Parallel Relay* utiliza duas câmeras, uma câmera grava um intérprete que realiza sinais baseando-se no script de tradução, e outra câmera grava os sinais de um segundo intérprete que repete os sinais realizados pelo primeiro intérprete (FELS et al., 2009).

Segala (2010) descreve a produção de vídeos em LS a partir de conteúdo textual direcionado ao público surdo, em situação de aprendizagem, nas seguintes etapas:

- a) produção do texto acadêmico para disponibilização do conteúdo on-line pelos designers instrucionais;
- b) leitura do texto acadêmico pelo tradutor;
- c) estudo do texto pelo tradutor da versão a ser disponibilizada on-line traduzida para a Libras;
- d) caso necessário, o tradutor seleciona palavras adequadas para organizar um glossário;
- e) faz-se a filmagem da tradução em Língua de Sinais;
- f) edição da filmagem;
- g) o próprio tradutor ou outra pessoa capacitada para a função faz a revisão da filmagem já editada;
- h) caso sejam detectados erros, a filmagem é refeita e reeditada;
- i) faz-se uma revisão final;
- j) põe-se o material à disposição (SEGALA, 2010, p. 37).

Em relação ao estúdio de gravação do intérprete, a NBR 15290 estabelece que este deve ter:

- a) espaço suficiente para que o intérprete não fique colado ao fundo, evitando desta forma o aparecimento de sombras;
- b) iluminação suficiente e adequada para que a câmera de vídeo possa captar, com qualidade, o intérprete e o fundo;

- c) câmara de vídeo apoiada ou fixada sobre tripé fixo;
- d) marcação no solo para delimitar o espaço de movimentação do intérprete (ABNT, 2008, p.9.)

Para uma boa visualização dos sinais, o aspecto do intérprete deve seguir as seguintes diretrizes:

- a) a vestimenta, a pele e o cabelo do intérprete devem ser contrastantes entre si e entre o fundo. Devem ser evitados fundo e vestimenta em tons próximos ao tom da pele do intérprete;
- b) na transmissão de telejornais e outros programas, com o intérprete da LIBRAS em cena, devem ser tomadas medidas para a boa visualização da LIBRAS;
- c) no recorte não devem ser incluídas ou sobrepostas quaisquer outras imagens. (ABNT, 2008, p.9.)

O conhecimento prévio pelo intérprete do material a ser traduzido é essencial para uma boa tradução, no caso os diálogos ou falas dos apresentadores/locutores da cena a ser interpretada. No caso de programas ao vivo, este conhecimento prévio não necessariamente será o *script* literal do que será dito, mas pode ser uma conversa a respeito do tema a ser apresentado, de modo a se formar um vocabulário e identificarem-se sinais que devam ser definidos.

Além do conhecimento prévio das falas a serem traduzidas, a uso de elementos, como mapas e diagramas apontados pelo locutor, contribuem na transmissão da mensagem do intérprete. Este uso de elementos adicionais pelo intérprete é apontado por Nascimento (2010), quando afirma que

a atuação do intérprete na esfera televisiva, no gênero jornalístico, é atravessada por elementos extralinguísticos que são fundamentais para o processo de interpretação (NASCIMENTO, 2010 p.7)

Um exemplo de integração do intérprete com o conteúdo audiovisual pode ser visto na figura 30.



Figura 30: totalidade do intérprete de Libras com o programa de TV
 Fonte: Nascimento (2010, p.6)

Para oferecer conteúdo em Libras na TV digital brasileira, Araujo et al. (2009) sugerem o uso de um método baseado na junção de unidades de vídeo pré-gravadas, armazenadas em um dicionário local, que são agrupadas de acordo com a tradução referente ao conteúdo que está sendo apresentado. Deste modo,

Cada sinal (ou expressão) pode ser representada por uma imagem animada ou um arquivo de vídeo [...] e tem um código (índice) associado com sua representação (ARAUJO et al., 2009, p.).

Com o objetivo de utilizar vídeos com LS, Krapež e Solina (1999) desenvolveram um sistema que permite a junção de diferentes tipos de videocliques com sinais de modo a formar sentenças em LS com transições suaves entre os diferentes sinais.

Santos, Oliveira e Oliveira (2010) descrevem um sistema para produção de legendas textuais que serve de base para a produção de vídeo em LS. O sistema é composto por um módulo que permite a criações e sincronização de legendagem textual e sua posterior tradução para Libras. Outro módulo do sistema realiza sua apresentação a partir da linguagem NCL (*Nested Context Language*). A figura 31 descreve a dinâmica para produção dos artefatos para acessibilidade seguido por Santos, Oliveira e Oliveira (2010).

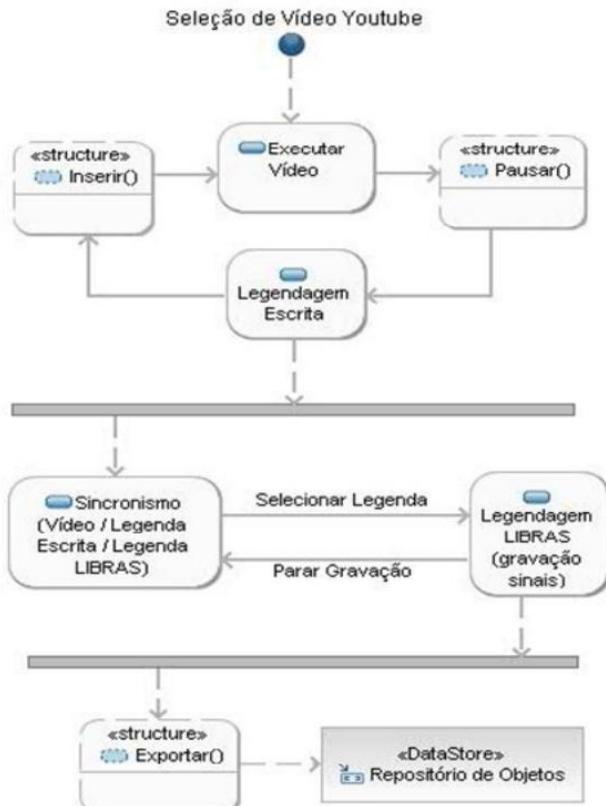


Figura 31: Diagrama de atividades para o processo para produção de legendas em vídeo.

Fonte: Santos, Oliveira e Oliveira (2010)

Um sistema similar é apresentado por Junior (2011), onde desenvolve um módulo para produção de legendas e outro para sua apresentação. O módulo de produção, o *SyncrLIBRAS*, é composto por um subsistema para produção de legendas escritas e outro para legendas em vídeo de LS. O módulo de apresentação, o *HiddenLIBRAS*, tem como principal vantagem a possibilidade de redimensionamento da janela de apresentação da legenda em LS.

De modo a facilitar a sincronização de legendas em LS com o audiovisual, Farhadi e Forsyth (2006) apresentam um método automatizado que realiza a sincronização a partir do tempo contido no *closed-captions*.

Para a produção de audiovisual legendado em LS, o guia para produção de mídia digital de JISC (2010) enumera algumas questões sobre o formato de vídeo a ser utilizado, como a relação entre tamanho de arquivo e qualidade visual e o posicionamento do intérprete na tela. Deste modo, um fator importante na apresentação da Libras é a qualidade da imagem apresentada, que é afetada quando o vídeo é codificado para reduzir a quantidade de dados utilizada tanto no seu armazenamento quanto na sua transmissão. Algumas técnicas de codificação e parâmetros associados são analisados no próximo subtópico.

3.2.1.1 Codificação de Vídeo em Línguas de Sinais

A codificação do vídeo em línguas de sinais deve permitir uma boa detecção visual de movimentos manuais e expressões faciais, ao mesmo tempo em que oferece redução no uso de dados necessários a seu armazenamento e transmissão, principalmente no cenário da televisão digital, conforme afirma Muir (2007):

Largura de banda é um recurso escasso na comunicação em canais de broadcast e existe a possibilidade de desenvolver esquemas de codificação altamente eficientes baseados no comportamento visual diante conteúdos sinalizados na televisão de modo a reduzir significativamente a quantidade de bits necessários à representação da informação visual (MUIR, 2007, p.214, tradução nossa)

Entretanto,

Para os surdos e os que tem dificuldade em ouvir, é importante que detalhes do movimento possam ser reproduzidos de modo que dedos, olhos e boca sejam distinguíveis mesmo quando os sinais são constituídos pelas duas mãos e braços em movimento com todos os dedos estendidos. Dedos borrados pelo movimento são toleráveis, embora dedos claramente visíveis são preferíveis (DEBEVC et al., 2009, p.280).

No sentido de melhorar a comunicação por vídeo em LS, Leelarasmee, Mahasith, Uthaichana (2005) apresentam um sistema para a expansão do vídeo apresentado na área destinada à LS, num modelo aplicado à TV analógica. Este sistema consegue duplicar a área do vídeo com uma qualidade aceitável pelo usuário surdo.

Muitos parâmetros podem afetar a clareza do vídeo, incluído o tamanho/resolução da imagem, taxa de quadros por segundo, quantidade de cores, uso de diferentes algoritmos de compressão de vídeo, taxas de transmissão variáveis e a forma de escaneamento dos quadros (*interlaced* ou *progressive scan*). A modificação destes parâmetros pode interferir com o processo de comunicação, por exemplo, maiores taxas

de quadros por segundo aumentam o entendimento da mensagem em LS (Hooper et al., 2007).

No contexto das LS,

o vídeo deve ser avaliado em termos de inteligibilidade da conversação, e não em termos de sua qualidade estética geral (Ciaramello; Ko; Hemami, 2010, p.71, tradução nossa).

Em termos técnicos a qualidade do vídeo para a conversação em LS pode ser medida pela qualidade individual de cada quadro do vídeo, pela taxa de quadros apresentada por segundo (fps), pelo atraso e pelo ruído (CHON et al., 2011).

Na intenção de reduzir a demanda de dados para a transmissão de vídeos em LS, Kountchev, Todorov, Kountcheva (2008) apresentam um algoritmo de extração de contornos de imagem voltado à comunicação por LS. Um quadro de vídeo com a aplicação do algoritmo é apresentado na figura 32.

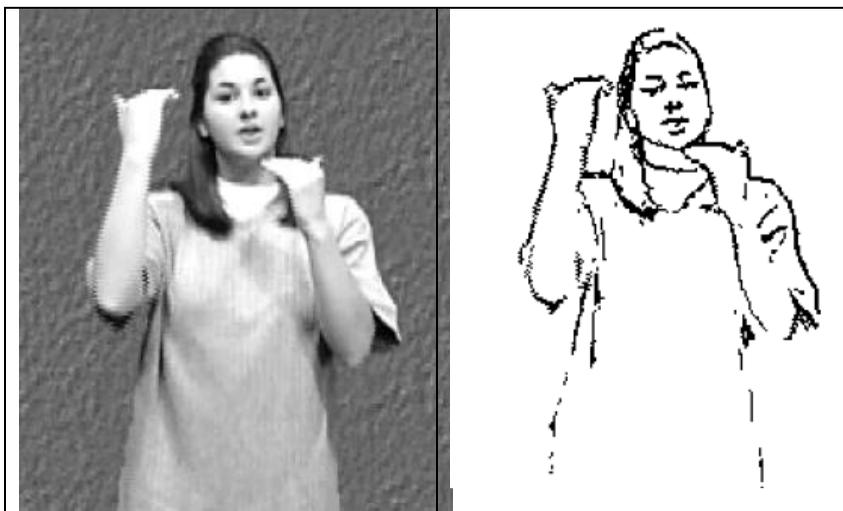


Figura 32: Quadro do vídeo original e quadro do vídeo com aplicação de algoritmo de extração de contornos.

Fonte : Kountchev, Todorov e Kountcheva (2008, p. 180)

Com a aplicação do algoritmo apresentado a compreensão dos sinais interpretados no vídeo em LS é mantida, e a taxa de compressão

de dados é elevada, permitindo um vídeo de 320 x 240 px a 15 fps com taxa de 23 kbps (KOUNTCHEV, TODOROV, KOUNTCHEVA, 2008).

As técnicas de compressão de vídeo consideram um modelo de percepção visual genérico, que permite que elementos essenciais à percepção dos gestos de uma língua de sinais sejam suprimidos no processo de codificação. De modo a evitar a perda de conteúdo significativo para comunicação em LS, Ciaramello, Ko e Hemami (2010) analisam a aceitação de duas formas de codificação de vídeo com LS, uma que prioriza a qualidade geral dos quadros e outra que prioriza a inteligibilidade das línguas de sinais, conforme a figura 33.

Os autores relatam que

Participantes com experiência significativa no uso de tecnologias de comunicação baseada em vídeo preferiram o vídeo codificado de acordo com o critério de qualidade, enquanto que usuários com pouca experiência preferiram o vídeo codificado de acordo com o critério de inteligibilidade (CIARAMELLO; KO; HEMAMI, 2010, p.78, tradução nossa).



Figura 33: comparação entre codificações de vídeo a 55kbps.

Fonte: Ciaramello, Ko, Hemami (2010, p.74).

Os autores constataram também que em velocidades acima de 80 kbps, o vídeo codificado de acordo com o critério de qualidade é melhor

aceito, pois sua inteligibilidade é suficiente. Entretanto em ambientes com velocidades menores foi preferida a codificação com priorização de conteúdo.

Em busca de uma codificação mais eficiente e realista, Muir e Richardson (2000) e Agrafiotis et al. (2004) realizaram experimentos de *eyetracking* para analisar os mecanismos de atenção do surdo em relação aos vídeos com LS e constataram que o surdo dedica a maior parte de sua resolução visual em regiões da face com o objetivo de capturar expressões faciais e formas da boca, enquanto que a visão periférica e de menor resolução é voltada a analisar movimentos mais rápidos e amplos, como os dos braços e mãos. Este comportamento demonstra a importância da região facial, como lábios, pálpebras e direção do olhar, na compreensão do significado dos sinais.

Aplicando estes conceitos, Muir (2007) apresenta uma otimização para a codificação de vídeos no formato h.264 de modo a melhorar a percepção dos conteúdos em línguas de sinais. O método prioriza a qualidade das áreas mais sensíveis à comunicação por LS e aumenta a performance de esquemas de codificação em termos de qualidade do vídeo percebida em ambientes com baixa taxa de bits (abaixo de 200kbps).

Seguindo a abordagem de Muir e Richardson (2005), Davies et al. (2007) conseguiram reduzir a taxa de bits em 30%, sem perda significativa de compreensão ou experiência de visualização, ao priorizar as regiões de maior importância na comunicação por LS em legendas abertas de sinais.

Nakazono, Nagashima e Ichikawa, (2006) obtiveram resultados satisfatórios na aceitação de vídeos em LS codificados com o algoritmo de codificação *AdaptSynch*, também baseados no trabalho de Muir e Richardson (2005).

O ITU-T *application profile serie H*, suplemento 1 (ITU-T, 1999) detalha os requisitos de qualidade para comunicação por meio de vídeos de LS, onde inclui como formato mínimo o CIF, com 352 x 288 pixel a uma taxa mínima de 25 quadros por segundo.

Na codificação de vídeo em LS nos formatos CIF (*Common Intermediate Format*), com resolução de 352 x 288 pixel x 25 fps, a melhor aceitação foi obtida com a taxa de 256 kbps. Já com o formato QCIF (*Quarter Common Intermediate Format*), que tem 176 x 144 pixel x 25 fps, a aceitação pode ser obtida com 128 kbps (NAKAZONO; NAGASHIMA; ICHIKAWA, 2006).

Entretanto o formato CIF não atende as especificações da norma ABNT no caso da transmissão em HD de 1920 x 1080, onde a área

destinada à janela de Libras deve ter $\frac{1}{4}$ da largura e metade da altura (480 x 540 pixel).

Outro fator inerente ao vídeo codificado é sua alta vulnerabilidade à perda de dados durante a transmissão, pois o algoritmo de codificação utiliza correlações temporais e espaciais entre os quadros, de modo que a perda de um dos pacotes transmitidos causa a perda de um quadro inteiro. Deste modo, em ambientes com alta taxa de perda de dados, a aplicação de métodos de rajadas de quadros intermediários, (I-frames e P-frames), torna o h.264 mais robusto (CHON et al., 2011).

Tran et al. (2010), visando aumentar a duração de baterias de dispositivos móveis utilizados na comunicação em língua de sinais, verificou a aceitação do uso de algoritmos com variação de resolução espacial (VSR - *variable spatial resolution*) e variação de taxa de frames (VFR - *variable frame rate*), ilustrados na figura 34. Em sua análise, a VFR causa a sensação de cortes/pausas no fluxo de vídeo (*choppy*), enquanto que a VSR causa a sensação de vídeo borrado (*blurry*). Quando as duas técnicas são usadas simultaneamente estes efeitos são minimizados consideravelmente, melhorando a qualidade de vídeo percebida pelo sinalizante.

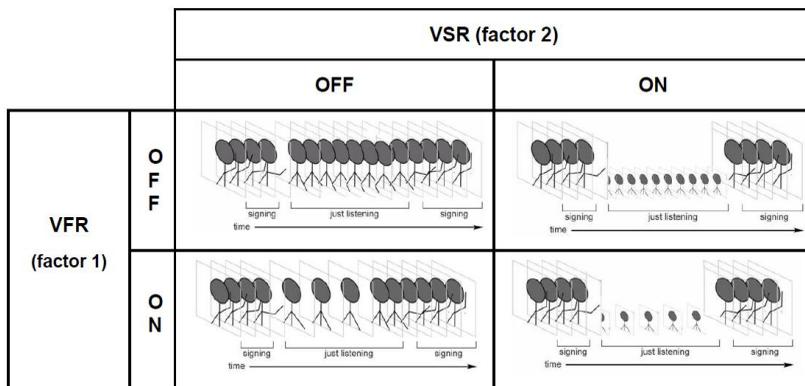


Figura 34: combinação de resolução espacial e taxa de quadros
Fonte: Tran et al. (2010, p. 119)

A codificação de vídeos em LS pode adotar técnicas que se diferenciam em termos de redução do tamanho de dados do vídeo e manutenção da qualidade de conversação. A redução de cores, como na codificação de Kountchev, Todorov e Kountcheva (2008), ou a variação de parâmetros de resolução de Tran et al. (2010), combinados com a

priorização de conteúdos (MUIR; RICHARDSON, 2005) são técnicas que se demonstram adequadas para aplicação na legendagem em LS.

Entretanto o uso de vídeo apresenta algumas restrições para as línguas de sinais:

- Os custos de produção de vídeo em padrões profissionais não é trivial;
- Existem problemas de continuidade. Fazer vídeos consistentes, ou seja, utilizando o mesmo sinalizante, com a mesma roupa e com o mesmo plano de fundo de modo que as frases sinalizadas possam ser unidas, complicam o processo de manutenção;
- Toda vez que algum detalhe do conteúdo é modificado, novos vídeos devem ser feitos, aumentando os custos;
- O armazenamento e transmissão de vídeos também podem ser problemáticos pois são arquivos grandes. Para usuários domésticos com conexões limitadas (dial-up), o tempo e custos envolvidos na transferência de sequências de vídeo podem ser proibitivos. Este aspecto pode se tornar menos significativo com o aumento de abrangência da banda-larga e melhorias na tecnologia de compressão de vídeo;
- Uma sofisticada arquitetura de servidor é necessária para a geração de conteúdo gerado ao vivo (KENNAWAY et al. 2007, p.17, tradução nossa).

Além disso,

Vídeos em línguas de sinais, apesar de sua popularidade atual, não são uma alternativa viável ao texto, por duas razões: primeiro eles não são anônimos – indivíduos realizando contribuições podem ser reconhecidos a partir do vídeo e portanto exclui aqueles que desejam manter sua identidade em segredo. Segundo, pessoas não podem editar facilmente ou adicionar conteúdo a um vídeo que alguém produziu (EFTHIMIOU et al., 2009, p.22).

Atualmente o uso do vídeo, meio essencial para comunicação em línguas de sinais devido a sua natureza visual/gestual, é possibilitado por recursos interoperáveis de captura e exibição e pelo aumento da capacidade de meios de transmissão nos ambientes digitais.

Apesar de sua popularidade, ele apresenta algumas restrições e necessita a consideração de fatores inerentes às línguas de sinais, tais como a priorização de conteúdo em algoritmos de codificação, o anonimato do interlocutor e a possibilidade de edição posterior.

Além disso, a junção de unidades de vídeo não confere realismo ao produto final, pois diante a grande variedade de combinações entre os sinais, é impossível ter todas as possíveis gravações com as mesmas características de iluminação e posicionamento do intérprete, ocasionando problemas de continuidade.

De modo a encontrar alternativas à mídia do vídeo tradicional, as próximas sessões analisam a possibilidade de legendas de LS na forma textual, o uso de glossários e as técnicas automatizadas, por exemplo a utilização de intérpretes virtuais para síntese de sinais.

3.3 LEGENDAS DE TEXTO EM LÍNGUA DE SINAIS

O armazenamento e processamento das línguas de sinais requer um formato para sua codificação. Deste modo, Bungeroth et al. (2006) apresentam um aplicativo para a anotação de língua de sinais em vídeos sobre previsão do tempo. Para realizar as anotações são utilizadas glosas, que consistem em representações semânticas de determinada língua de sinais e formam conjuntos de termos referentes a um domínio específico.

As glosas indicam quais sinais devem ser realizados por meio de palavras da língua oral escrita e podem utilizar marcadores adicionais. A interface de um aplicativo que permite o uso de glosas pode ser visto na figura 35.

The screenshot shows the ELAN software interface. The main window displays a video of a weather forecast with a sign language interpreter. The interface includes a menu bar (File, Edit, Search, View, Options, Help) and a toolbar. A 'GLOSSE' window is open, showing a list of terms and their corresponding sign language glosses. The timeline window at the bottom shows the following data:

Time	GLOSSE	Wortart	Gebärdensatz	Deutscher Satz	Deutsch
00:16:01.000	FREUNDLICH	ADJ			nieder freundliches Frühlingwetter, dann macht es Tiefdruckgebieten.
00:16:02.000	WETTER	N			
00:16:03.000	TIEF-emp_GEBIET	N			
00:16:04.000	NOCH-konz	ADV			Noch fließt südliche Warmluft nach Deutschland, aber ab Sonntag lenkt d
00:16:05.000	SÜDE	NH			
00:16:06.000	WARM	ADJ			
00:16:07.000					

Figura 35: Aplicativo ELAN utilizando glosas para transcrição do conteúdo em língua de sinais.

Fonte: Bungeroth et al.(2006, p.3)

Glosas podem ser usadas para transmitir mensagens em situações em que o meio de transmissão é restrito e não pode se enviar um vídeo (KAMATA, FUJII, WATANABE, 2009), entretanto por estarem baseadas

nas línguas orais, não permitem expressar totalmente as características das LS.

Os sistemas de escritas de sinais demonstram-se mais adequados para representar as LS na forma escrita. Eles servem como modelos sobre como os movimentos podem ser estruturados e armazenados em um ambiente computacional (MAURER; STUBENRAUCH; CAMHY, 2003).

A escrita de sinais pode adotar diferentes sistemas de notação, como *HamNoSys*, *Elis* ou *SignWriting*. Diante esta variedade de opções, o sistema de escrita *SignWriting* demonstra ser adequado ao contexto da comunicação escrita em LS (SILVA, 2009).

ao ler em *SignWriting*, o leitor é capaz de associar informações já adquiridas às novas informações. Acrescentar, interpretar, resumir, tudo é possível nesta troca de informações entre texto e leitor na língua de sinais (SILVA, 2009, p.99).

A escrita de sinais *SignWriting* já vem sendo utilizada em menus de aplicativos desktop, web e em mensagens em dispositivos móveis (PONTES; ORTH, 1999; AHMED; SEONG, 2006). A figura 36 demonstra a pré-visualização de mensagem em escrita em sinais usando o sistema *SignWriting*.

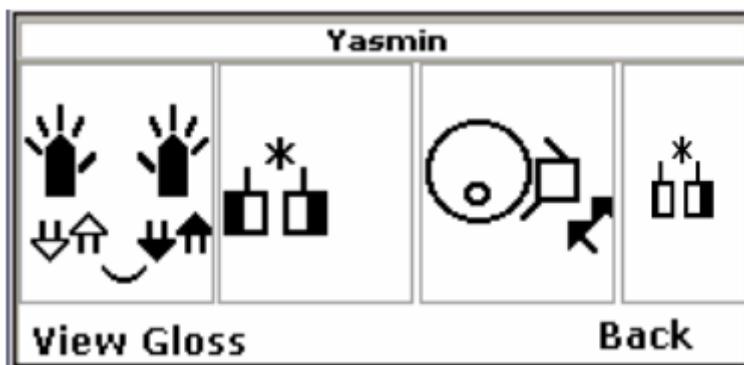


Figura 36: Visualização de mensagem em escrita de sinais em dispositivo móvel
Fonte: Ahmed e Soung (2006, p.4)

Ao encontrar um índice de aceitação de 88% em relação ao uso de escrita de sinais na comunicação em celulares por meio de

mensagens, Ahmed e Seong (2006) afirmam que devido ao fato de celulares terem uma grande prevalência na comunidade surda, estes, ao incorporarem alfabetos para a escrita de sinais, podem mudar significativamente o modo de comunicação dos surdos.

Os próprios usuários surdos afirmam que o uso da escrita de sinais em ambientes de comunicação pode facilitar a comunicação entre pares (SOUZA; PINTO, 2003). A figura 37 apresenta a opinião de surdos a respeito do uso de escritas de sinais para comunicação em ambiente digitais.



Figura 37: Opinião de surdos sobre uso de software de comunicação por escrita de sinais.

Fonte: Souza e Pinto (2003).

Os aplicativos de comunicação que utilizam escrita de sinais oferecem

um novo meio/recurso para a escrita da Libras e para comunicação a distância entre surdos e entre surdos e ouvintes. Além disso, esta ferramenta pode servir como estímulo à aquisição da língua oral escrita pelos surdos, de maneira a facilitar a estes o acesso à informação, à cultura, ao intercâmbio e ao trabalho (SOUZA; PINTO, 2003, p.9).

Um exemplo de escrita de sinais na web pode ser visto na figura 38 a seguir, onde é utilizada no blog de um surdo. Nota-se seu papel de marcador cultural e meio para afirmação cultural (ZAPPE, 2010).

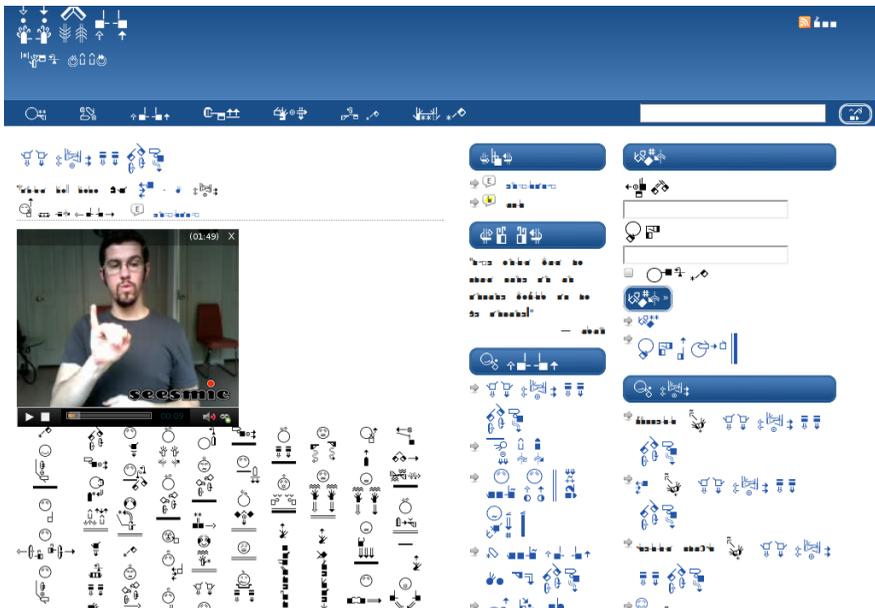


Figura 38: Web site em escrita de sinais.

Fonte: <http://frostvillage.com/blog/lang/ase/>

Diante estas aplicações, busca-se integrar a escrita de sinais nos processos para produção de conteúdo audiovisual acessível aos surdos.

A figura 39 apresenta um esquema de visualização de legendas em escrita de sinais no contexto do audiovisual.



Figura 39: Proposta de aplicação de legenda com escrita de sinais.

Seu uso em ambientes digitais pode ser constatado diante as iniciativas apresentadas nas sessões anteriores. A aplicação de escrita de sinais na legendagem do audiovisual baseia-se na ideia de que esta modalidade de representação da língua permite seu entendimento (SEGALA, 2010).

O próximo tópico analisa algumas alternativas relacionadas à produção de conteúdo com escrita de sinais.

3.3.1 Produção de Legenda em Escrita de Sinais

Não foram encontradas pesquisas diretamente relacionadas à produção de legendas com escrita de sinais. O *SW-Captioner* é um projeto de produção de legendas que utiliza escrita de sinais (BARTH, 2008), entretanto não foi encontradas maiores informações sobre o projeto.

Deste modo são descritas as iniciativas relacionadas aos editores de texto desenvolvidos para esse tipo de escrita (SUTTON; GLEAVES, 1995) e que podem apoiar a produção deste tipo de legenda e a descrição de um formato de arquivos para seu armazenamento e distribuição.

Um editor de texto em escrita de sinais, denominado *Delegs*, pode ser visto na figura 40, onde o usuário digita um texto em língua de sinais germânica e a tradução correspondente é exibida em escrita de sinais. A intenção deste editor é permitir que surdos aprendam a língua alemã a partir da referência em escrita de sinais da *Deutsch GebärdenSchirft* (DGS).

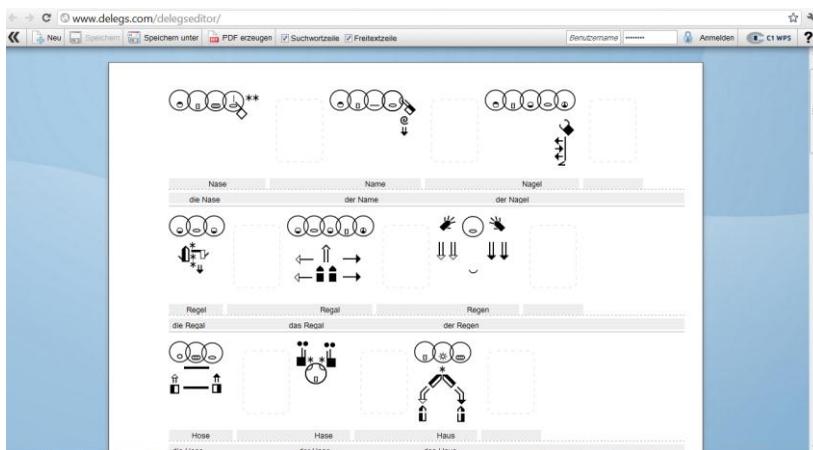


Figura 40: Editor de texto em escrita de sinais

Fonte: <http://www.delegs.com/DelegsPage/>

Este editor aceita como entrada o texto em língua oral, porém não permite a edição direta da escrita em sinais.

O *SignWriter Studio* (SIGN WRITER STUDIO, 2012) é um editor de texto de sinais que permite a manipulação direta dos símbolos

dos sinais. Outro editor de texto que segue o mesmo princípio é o *SignText* (SIGNTEXT, 2012), que pode ser visto na figura 41.

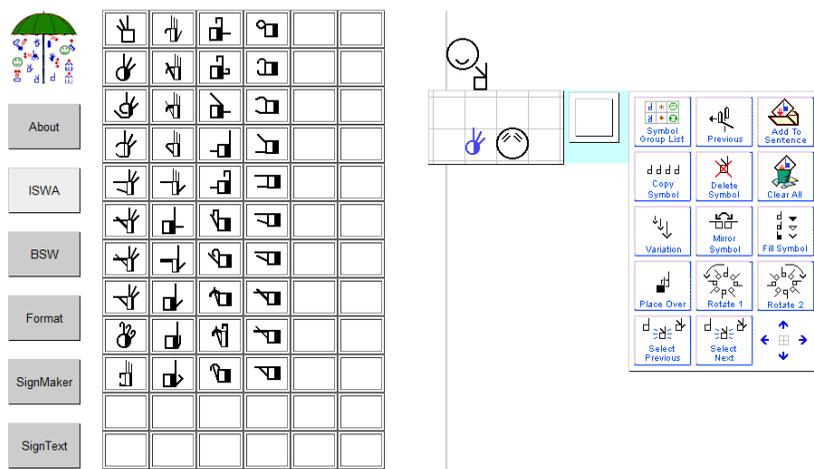


Figura 41: Editor de texto de sinais SignText

Fonte: SignText (2012).

O *SW-Edit* (SW-EDIT, 2012) é outro editor de texto de sinais, produzido pela PUC de Pelotas, tem sido utilizado pelos alunos do curso Letras-Libras da UFSC e pela comunidade brasileira de escrita de sinais. Sua interface pode ser vista na figura 42. Nota-se a diferença de disposição dos símbolos da escrita de sinais entre os aplicativos *SW-Edit* e *SignText*.

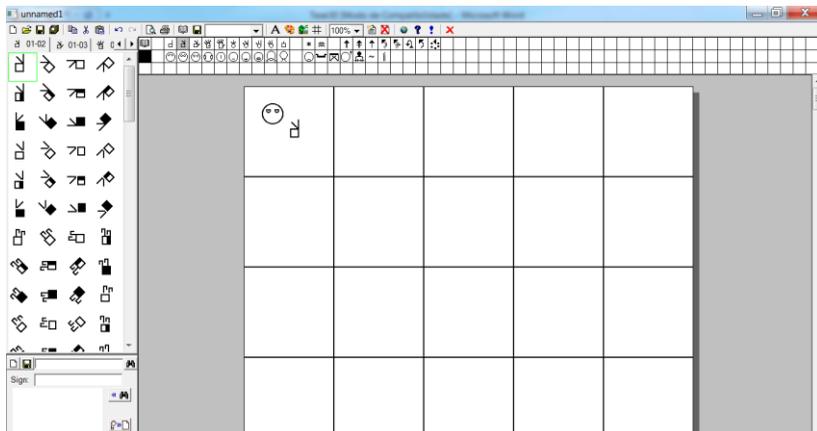


Figura 42: Editor de texto de sinais SW-Edit.

Fonte: SW-Edit (2012).

O *SignText*, *SW-Editor* e *SignWriter Studio* são editores que permitem a entrada direta de caracteres correspondentes aos constituintes fonológicos das LS, enquanto que o *Delegs* utiliza o texto da língua oral para inserção de conteúdo.

Souza (2005b) descreve uma arquitetura de software para serviços de oferta de escrita em línguas de sinais na web.

Um fator importante para a adoção de escrita de sinais é devido às estratégias para a tradução automatizada das LS, as quais necessitam de uma codificação dos sinais baseada numa representação escrita.

A inteligibilidade de mensagens sintetizadas a partir de *SignWriting* se mantém a taxa de reconhecimento de 82%, com uma precisão de 66% (MOEMEDI; CONNAN, 2010).

Um requisito para uso de escrita de sinais para acessibilidade do audiovisual é sua associação com informações temporais para o sincronismo de legendas. Para este fim a SWML é estendida para representar informação temporal associada à escrita de sinais. Os detalhes desta extensão são apresentados na aplicação do modelo de referência.

Necessita-se também a definição da quantidade de caracteres que podem ser exibidas por unidade de tempo para que a leitura seja adequada, de modo análogo à forma escrita nas línguas orais.

A produção de conteúdo em escrita de sinais está associada aos glossários de LS, que permitem uma melhor compreensão das legendas e são descritos na sessão seguinte.

3.4 GLOSSÁRIOS DE LÍNGUAS DE SINAIS

No contexto do audiovisual acessível um glossário, ou dicionário, serve como elemento de apoio tanto à produção de legendas, fornecendo um corpo de vocabulário, quanto ao entendimento por parte do usuário de termos desconhecidos presentes na tradução.

Diante uma perspectiva de uso de legendas em texto na língua oral e de legendas em LS, um glossário bilíngue promove a integração destes dois espaços conceituais.

As figura 43 a 45 apresentam um exemplo de dicionário em Libras /português descritos por Souza (2005b).

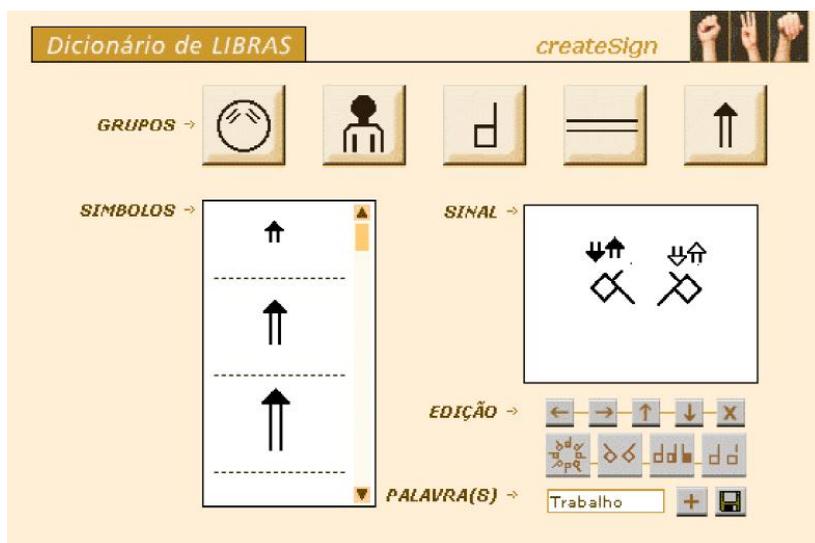


Figura 43: Interface de cadastro de sinal em dicionário.

Fonte: Souza (2005b, p.70).

A figura 43 descreve a tela para cadastro de determinado termo no dicionário, no caso o cadastro da palavra “trabalho”.

Já as figuras 44 e 45 descrevem a busca de termos a partir dos paramentos fonológicos da língua de sinais (figura 44), ou a partir dos caracteres da escrita da língua oral (figura 45).

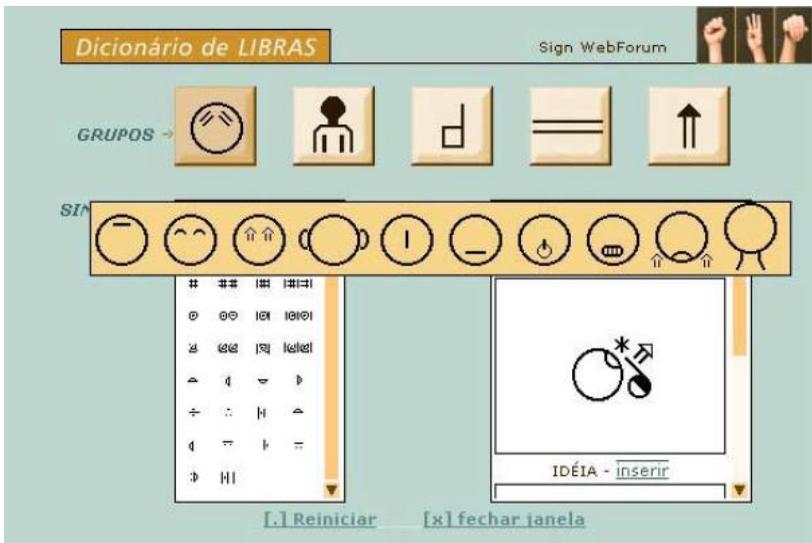


Figura 44: Interface de consulta em dicionário a partir de escrita de sinais.
 Fonte: Souza (2005b, p.105).



Figura 45: Interface de busca em dicionário a partir de escrita oral.
 Fonte: Souza (2005b, p.105).

Um exemplo de glossário é apresentado por Oliveira (2010), onde demonstra questões relacionados a um glossário utilizado no curso de licenciatura e bacharelado em Letras-Libras. O glossário pode ser visto na figura 46 que apresenta a descrição do termo “Ambiente Virtual”, no contexto do curso de licenciatura e bacharelado em Letras-Libras da UFSC.

Ambiente Virtual:



"Espaço na internet onde funciona o curso online, neste caso, Curso Letras LIBRAS"

Figura 46: Descrição do termo "Ambiente Virtual" no glossário Letras-Libras.
Fonte: www.libras.ufsc.br

Johnston e Napier (2010) descrevem o processo de desenvolvimento e a dinâmica de uso de um dicionário de sinais relativos a informações médicas, apresentado na figura 47, onde

Usuários podem acessar uma base de dados de sinais em Auslan, ver o vídeo digitalizado correspondente a cada sinal, e navegar pela base de dados recuperando informações sobre a origem de cada sinal, seu significado, uso e gramática. Podem oferecer *feedback* sobre sinais existentes (corrigindo ou confirmando sua forma, uso e significado), registrando sinais de sua comunidade que estão faltando no dicionário e também sugerir neologismos de sinais que estão utilizando em

conjunto com seus amigos ou intérpretes (JOHNSTON; NAPIER, 2010, p. 266, tradução nossa)

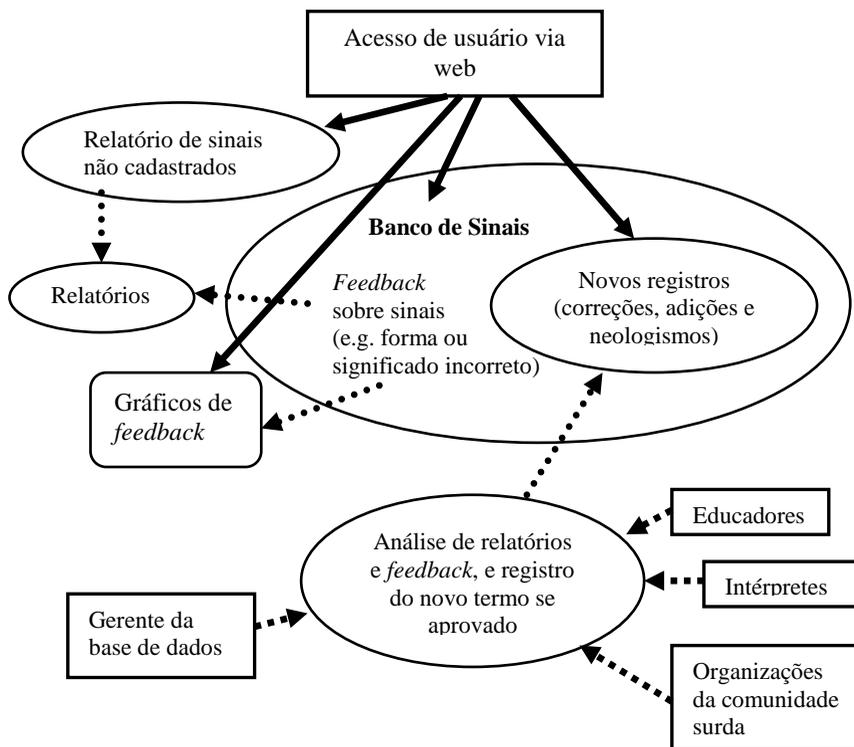


Figura 47: dinâmica de desenvolvimento e uso de glossário em LS
Fonte: Traduzido de Johnston, Napier (2010, p.266)

Uma questão inerente as LS é a variação dos sinais entre comunidades, mesmo na Libras, determinado conceito pode ter diferentes representações gestuais. Para resolver o problema de localidade dos sinais, cada comunidade deve ter acesso a um dicionário próprio, além da possibilidade de dicionários compartilhados entre comunidades.

O desenvolvimento de um dicionário específico da área de filosofia é apresentado por Rocha (2008). Nesse dicionário buscou-se

evitar a formação de dialetos regionais a partir de sinais provisórios, e a consequente restrição de comunicação entre usuários da língua de sinais.

Já no sentido de permitir a variação regional, Jemni e Elghoul (2008) apresentam um sistema onde cada comunidade pode adicionar e editar seus próprios sinais, que são validados por um administrador.

Para a busca de termos em LS no dicionário são utilizados parâmetros como o de configuração de mão ou ponto de contato. A figura 48 apresenta uma interface de busca de sinais a partir destes parâmetros.

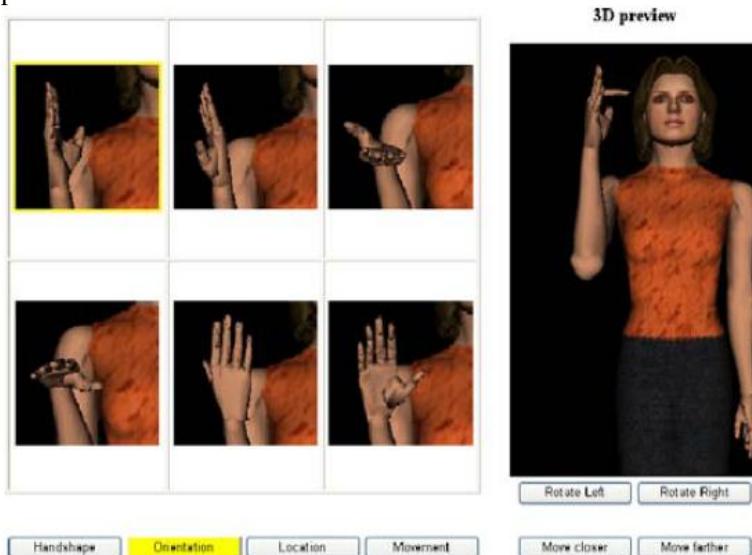


Figura 48: interface de busca de sinais a partir de cheremas representados por avatar

Fonte: Buttussi, Chittaro, Coppo, (2007, p.67).

Esquemas de busca em glossário de LS baseados em ontologias visuais são descritos por Gennari e Mascio (2007) e Felice, Mascio e Gennari (2008) e Tania, Mauro e Rosella (2008).

O esquema de Felice, Mascio e Gennari (2008) é utilizado no “Dizionario Elettronico di Base Bilingue Lingua Italiana dei Segni-Italiano”(DIZIONARIO, 2012) , apresentado na figura 49.



Figura 49: Interface com parâmetros de busca em LS.
Fonte: Dizionario (2012).

A figura 50 apresenta o resultado da busca ao sinal correspondente ao termo “lavoro”. São apresentados sua definição em vídeo com LS, exemplos de aplicação, informações de uso, variações regionais e outros possíveis significados para o termo buscado.



Figura 50: Resultado da busca por sinais.

Fonte: Dicionario (2012).

O SignBank (SignBank, 2012) , visto na figura 51, é um banco internacional de dicionários de sinais.

Figura 51: Lista de dicionários em língua de sinais SignBank.

Fonte: <http://www.signbank.org/signpuddle/>.

Cada dicionário utiliza uma base de sinais por meio do *SignPuddle*, o qual permite a busca de termos por palavras da língua oral ou por símbolos de escrita de sinais. A figura 52 descreve o resultado da busca por sinais no *SignPuddle*. Este dicionário é alimentado por meio do editor Sign Text apresentado anteriormente.

SignPuddle Online v2.0
 Pesquisar por palavras

Home

Dicionário Brasil

Pesquisar a linguagem falada com Resultados do Palavras

Termos e Títulos	<input type="text" value="amargo"/>
	<input checked="" type="radio"/> Qualquer parte da palavra <input type="radio"/> Início da palavra <input type="radio"/> Palavra exata
Texto	<input type="text"/>
Fonte	<input type="text"/>
Puddle página	<input type="text"/> Ver todas as entradas

amargo

Source: www.librasescrita.com.br
 Search location: [exact](#) or [approximate](#)
 Search symbols: [exact](#) or [base](#)
 Modificado: April 04, 2012 16:39
 Puddle página 2845

[Customize Sign](#) [Copiar sinal](#)

Figura 52: Pesquisa por palavra em dicionário de LS.
 Fonte: <http://www.signbank.org/signpuddle2.0/>.

Os glossários de LS podem apoiar os sistemas de tradução automática, que são analisados na sessão seguinte.

3.5 SISTEMAS DE TRADUÇÃO AUTOMÁTICA

Os sistemas de tradução automática podem ser analisados diante os tipos de tradução definidos por Jakobson (1995): a intralingüística, a interlingüística e a intersemiótica. A classificação ocorre em função da língua e do tipo de mídia de entrada e de saída, sendo as traduções entre o texto, a voz e o vídeo em LS os meios mais significativos.

A tradução do português oral para as LS consiste numa tradução interlingüística, por se tratar da conversão da mensagem de uma língua para ser apresentada em outra, e também intersemiótica, pois os meios utilizados para a apresentação dessa mensagem são diferentes: oral/auditivo e gestual/visual.

Os sistemas de tradução automática fundamentam-se em três ideias principais, conforme enumera Segala (2010):

- a) a ideia de que os sistemas de tradução automática intermodal acompanham, em linhas gerais, os princípios, abordagens e técnicas já desenvolvidas para os sistemas intramodais (de uma língua oral auditiva para outra língua oral-auditiva);
- b) a ideia de que os sistemas de tradução intermodal se subdividem, na verdade, em dois subsistemas: o de tradução de uma língua oral-auditiva para um sistema de escrita da língua gestual-visual; e o de síntese de sinais (gestual-visuais) a partir desse sistema de escrita; e
- c) a ideia de que a complexidade da tarefa está evidentemente relacionada ao sistema de escrita da língua gestual-visual adotado (SEGALA, 2010, p. 27).

Relacionado ao resultado da tradução das LS, as técnicas para sua apresentação podem ser divididas em articulatórias, de captura de movimento e de vídeos pré-gravados, e diferem entre si em termos de realismo, facilidade de edição e escalabilidade para a síntese de línguas de sinais (MALDONADO, 2007).

As técnicas articulatórias fazem uso de parâmetros que orientam a configuração das articulações do esqueleto de agentes virtuais; o uso de captura de movimento adota avatares articulados, controlando-se o

movimento do esqueleto do avatar e por consequência de pontos na malha do avatar. Os vídeos pré-gravados, por sua vez, não utilizam avatares, mas sim filmagens realizadas com intérpretes reais.

Uma comparação apresentada no quadro 7, indica que os vídeos pré-gravados oferecem maior realismo ao conteúdo, entretanto não possuem boa escalabilidade. Em um lado oposto, o uso de técnicas articulatórias em avatares oferece maior escalabilidade, apesar da perda de realismo.

Quadro 7: comparativo de técnicas para a síntese de conteúdo em línguas de sinais.

Fonte: Maldonado (2007, p.14).

Tipo	Realismo	Facilidade de edição	Escalabilidade
Articulatórios	+	+++	+++
Captura de movimentos	++	++	++
Vídeos pré-gravados	+++	+	+

As próximas sessões descrevem abordagens para a produção de conteúdo em LS, como o reconhecimento e tradução de línguas de sinais a partir de vídeo, a concatenação de vídeos a partir de voz, e a geração de LS a partir de textos por meio de técnicas de animação.

3.5.1 Reconhecimento das Línguas de Sinais a partir de vídeo em LS

Esta sessão trata da tradução intersemiótica, que não engloba a tradução entre línguas, mas sim entre os meios de representação de vídeo para escrita. Este tipo de tradução pode ser observada em tecnologias de reconhecimento de voz, no caso de línguas orais, ou de reconhecimento de gestos, no caso de línguas de sinais (WANG et al., 2008; ALLEN; ASSELIN; FOULDS, 2003; HONG; TURK; HUANG, 2000).

Dreuw et al. (2010) desenvolvem um projeto que visa trabalhar com o reconhecimento e tradução das línguas de sinais de modo análogo ao reconhecimento de fala e tradução estatística por máquina das línguas faladas. O objetivo da iniciativa é traduzir a língua de sinais de modo contínuo a partir da aplicação de técnicas de visão computacional em seu reconhecimento, permitindo a oferta de recursos de tradução de vídeo para texto em LS.

A abordagem de Dreuw et al. (2010) visa superar os problemas citados por Ong e Ranganath (2005), em relação ao reconhecimento automático de LS a partir de vídeos,

A análise automática de gestos em LS já percorreu um longo caminho desde seus primórdios quando meramente classificava sinais estáticos e alfabetos. Trabalhos atuais podem lidar com sinais dinâmicos que envolvem movimentos e que aparecem em sequências contínuas. Muita atenção tem sido focada na construção de sistemas de reconhecimento de grandes vocabulários. A esse respeito, sistemas baseados em visão computacional ficam atrás de sistemas que adquirem dados dos gestos com dispositivos de medição direta. A robustez do ambiente de captura de imagem também é um problema. Dois aspectos do reconhecimento de gestos que não receberam muita atenção é a construção de sistemas de reconhecimento independentes de sinalizantes e a abordagem de aspectos mais difíceis da sinalização, tais como inflexões gramaticais e sinais miméticos. Além disso, os sinais não manuais tem recebido pouca atenção. O entendimento dos sinais não manuais e sua interpretação em conjunto com o reconhecimento de gestos são vitais para o entendimento da

comunicação em LS (ONG, RANGANATH, 2005, p. 886, tradução nossa).

O sistema de reconhecimento ao utilizar a tradução estatística requer um grande *corpora* de anotações bilíngues. Além disso, no reconhecimento da LS devem ser tratadas as variações intrapessoais, como a velocidade e aparência do sinal, assim como as diferenças interpessoais presentes na apresentação das LS (DREUW et al., 2008).

No contexto nacional, o Sensor Libras é um sistema para tradução automática da Libras para o português que utiliza luvas de dados (TAVARES; BARBOSA; LEITHARDT, 2009), entretanto está focado apenas nos aspectos manuais dos sinais.

Para o reconhecimento e tradução das línguas de sinais de modo contínuo é necessário seu melhor entendimento linguístico, pois o reconhecimento de grandes vocabulários em línguas de sinais a partir de modelos de palavras inteiras não é adequado, sendo inviável produzir material suficiente para treinar os modelos de cada palavra (DREUW et al., 2010).

Como solução para o reconhecimento de grandes vocabulários por meio de visão computacional, Dreuw et al. (2010) propõem o uso de modelos fonológicos que dividam cada palavra em subunidades. Desta forma o treinamento pode focar no reconhecimento de cada uma dessas subunidades, e as palavras podem ser modeladas como a junção destas subunidades, fornecendo um modelo mais robusto para o reconhecimento e permitindo que palavras que não estiverem no conjunto de treinamento também possam ser reconhecidas.

Outros aspectos devem ser considerados no reconhecimento automático de LS:

- Simultaneidade: a maior diferença entre o reconhecimento de línguas faladas diante as línguas de sinais é o uso em paralelo de diferentes canais de comunicação, como expressões faciais, movimentos da mão e postura corporal.
- Espaço de sinalização: entidades, tais como pessoas e objetos, pode ser armazenadas no espaço em torno do sinalizante por meio da execução do sinal em determinada localização e depois apenas referenciá-las apontando para a localização;

- Coarticulação e Epêntese: assim como no reconhecimento da fala, o reconhecimento de sinais deve considerar os efeitos da coarticulação e da epêntese. A epêntese consiste nos movimentos que ocorrem regularmente quando se move de um estado final de um sinal para o início de outro. Este movimento não tem significado em si mas contribui para a informação fonética que é percebida;
- Silêncio: no reconhecimento de fala espaços de silêncio podem ser detectados pela medição da energia do sinal de áudio dentre as sentenças. Entretanto o silêncio na língua de sinais não pode ser reconhecido pela ausência de movimento no vídeo, pois existem sinais que são executados com a parada temporária em uma pose particular no espaço de sinalização durante determinado tempo. Além disso a posição de descanso das mãos podem ser em outro lugar no espaço de sinalização.
- Dialeto e diferentes intérpretes: o reconhecimento deve ser robusto para que possa ser independente da pessoa que o opera e deve tratar diferentes dialetos de línguas de sinais. (DREUW et al., 2008, tradução nossa).

No processo de reconhecimento são utilizados técnicas como o rastreamento visual de objetos e extração de parâmetros, os quais são fornecidos a um software de classificação do objeto rastreado. Devido à imprecisão no reconhecimento por visão computacional, Aznar, Dalle, Ballabriga (2006) ao adotar a abordagem *bitewise* para codificação dos sinais, afirmam que

no reconhecimento de vídeo, a abordagem *bitewise* oferece a vantagem da finalização *fuzzy*: caso um símbolo específico não seja completamente reconhecido, pode-se definir os bits para identificar quais parâmetros foram reconhecidos (como a rotação, elemento, etc.) e os

outros parâmetros podem ser informados pelo usuário, ou deduzidos de acordo com o contexto (sinais previamente usados, etc.)(AZNAR; DALLE; BALLABRIGA, 2006, p.31).

É possível o treinamento de um sistema de reconhecimento de LS a partir de vídeos que possuem legendas textuais e janela com o intérprete de Libras (COOPER; BOWDEN, 2009).

De modo similar, Buehler, Everingham e Zisserman (2009) propõem um modelo baseado em aprendizagem de máquina com supervisionamento fraco, que pode aprender sinais da BSL (*British Sign Language*) a partir de transmissões televisivas que contenham tanto legendas textuais e janela de intérprete de sinais. Para a aprendizagem utilizam técnicas de visão computacional com rastreamento do tronco e descritores extraídos das mãos.

Souza e Pistori (2005) desenvolvem um *plugin* para o *ImageJ* que consiste em um extrator de características baseado em momentos da imagem. O *plugin* é capaz de extrair parâmetros de imagens capturadas, os quais servem de dados de entrada para o processo de reconhecimento de sinais (SOUZA; PISTORI, 2005; PISTORI, 2006).

O reconhecimento da língua de sinais deve ocorrer de modo contínuo e gerar uma representação linguística, como a escrita de sinais. Esta representação dos sinais pode gerar uma animação gráfica a partir da composição dos componentes fonológicos dos sinais (EFTHIMIOU et al., 2009).

A próxima sessão descreve técnicas e procedimentos para o reconhecimento de voz e geração de LS a partir de texto e vídeos pré-gravados.

3.5.2 Geração de LS a partir de Vídeos Pré-gravados

Este tópico aborda o uso de vídeos pré-gravados para a apresentação de conteúdo acessível em LS. As abordagens para geração de conteúdo em LS com a concatenação de vídeos podem ser orientadas pela tradução de textos da língua oral para a LS ou auxiliadas por técnicas de reconhecimento de voz.

Um sistema de reconhecimento de voz que gera um vídeo em língua de sinais é apresentado por Oi, Tang e Wai (2009). Este sistema traduz sons de palavras na língua inglesa e apresenta vídeos na língua de sinais da Malásia. O processo principal do sistema é apresentado na figura 53. Entretanto o nível de tradução engloba apenas palavras, não contemplando aspectos gramaticais com a adaptação das estruturas das frases.

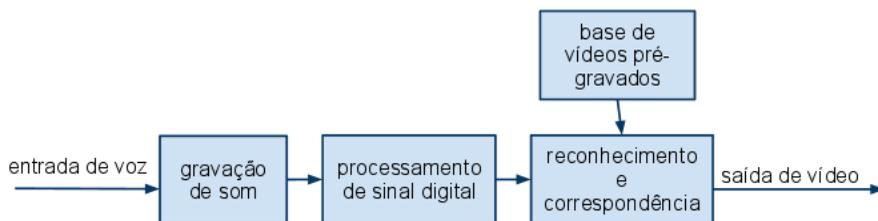


Figura 53: arquitetura de sistema de reconhecimento de voz para geração de vídeo em língua de sinais.

Fonte: Adaptado de Oi, Tang, Wai (2009, p 871)

O alinhamento temporal entre o conteúdo em LS e o texto de uma língua oral pode ser realizado pelo “uso de um modelo discriminativo de palavras” (FARHADI; FORSYTH, 2006, p.1476, tradução nossa).

San-Segundo et al. (2008) apresentam um sistema de tradução do espanhol para a língua de sinais espanhola que utiliza o reconhecimento de fala, obtendo uma taxa de erro de sinais de 31,6%.

A tradução do texto para geração de LS parece ser a alternativa mais viável para tornar acessível o conteúdo da legenda em português para o surdo que prefere a LS. Entretanto esta tradução exige que sejam considerados fatores que eliminam sua simplicidade, pois

Existem máquinas de tradução que utilizam a abordagem de converter diretamente a língua escrita para sinais, ignorando aspectos linguísticos específicos e consequentemente gerando conteúdos inacessíveis a audiências surdas que

possuem dificuldade na compreensão de textos escritos (HUENERFAUTH, 2005).

De modo a evitar esta conversão direta de texto para sinais, o FALibras MT é um ambiente para a autoria de tradutores automáticos do português para a Libras que se baseia em uma memória de tradução definida por um dicionário de símbolos e um dicionário de vídeos que estão associados a um domínio de conhecimento ou contexto (TAVARES; CORADINE; BREDA; 2005), conforme figura 54.

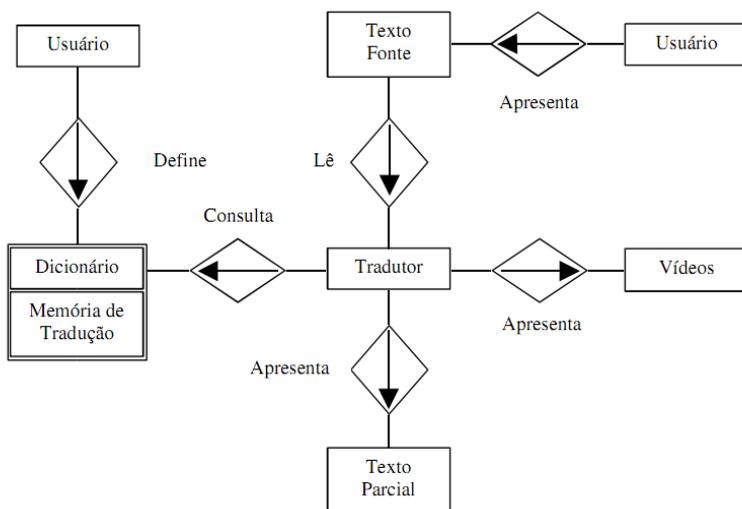


Figura 54: Estrutura Funcional do FALibras
Fonte: Tavares, Coradine e Breda, (2005, p.2103)

A arquitetura do ambiente é apresentada na figura 55, onde um texto fonte em português é traduzido em estruturas sentenciais de Libras que em seguida são traduzidos em vídeos. O dicionário de símbolos armazena estruturas sentenciais do português e sua correspondência com estruturas sentenciais de Libras. O dicionário de vídeos contém as formas gestuais correspondentes às estruturas sentenciais em Libras.

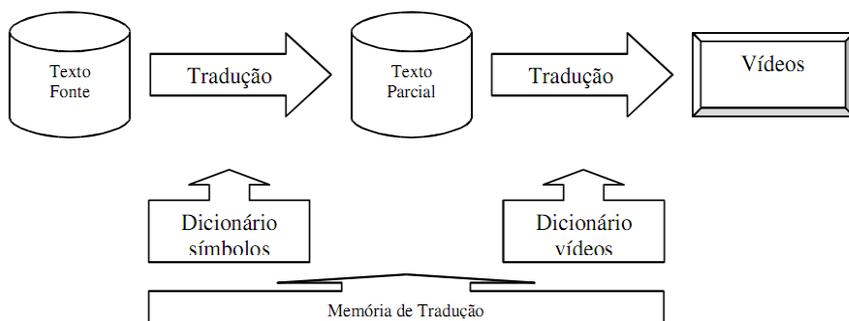


Figura 55: Arquitetura do módulo de tradução do FALibras MT
 Fonte: Tavares, Coradine, Breda (2005, p.2102)

A tradução do texto para vídeos de LS demonstra potencial para aplicação no contexto do acesso do audiovisual, porém depende da existência de uma legenda em texto ou do reconhecimento de áudio. As iniciativas apresentadas nessa sessão se baseiam na junção de vídeos pré-gravados, fato que diminui a escalabilidade e qualidade da legenda final.

A síntese de vídeo em LS com agentes virtuais a partir de sua representação escrita demonstra maior escalabilidade e é apresentada no tópico seguinte.

3.5.3 Síntese de Línguas de Sinais com Intérpretes Virtuais

A pesquisa em interfaces de comunicação tem focado no projeto de agentes conversacionais que são visíveis em interfaces, tais como faces animadas falantes, exibindo expressões faciais, síntese de fala com sincronização labial, e gráficos tridimensionais com movimentos de corpo complexos (BALDASSARI; CERESO; SERÓN, 2008).

Nesse sentido, o desenvolvimento de artefatos para acessibilidade ao audiovisual pelos surdos pode utilizar as tecnologias de avatares (GHOUL; JEMNI, 2010; SHOHIEB et al., 2009; RIEGER; BRAUN, 2003).

Avatares podem ser utilizados por exemplo no apoio ao desenvolvimento de estratégias de leitura labial em crianças com surdez profunda, conforme exemplificado por Beskow (2003) na figura 56.

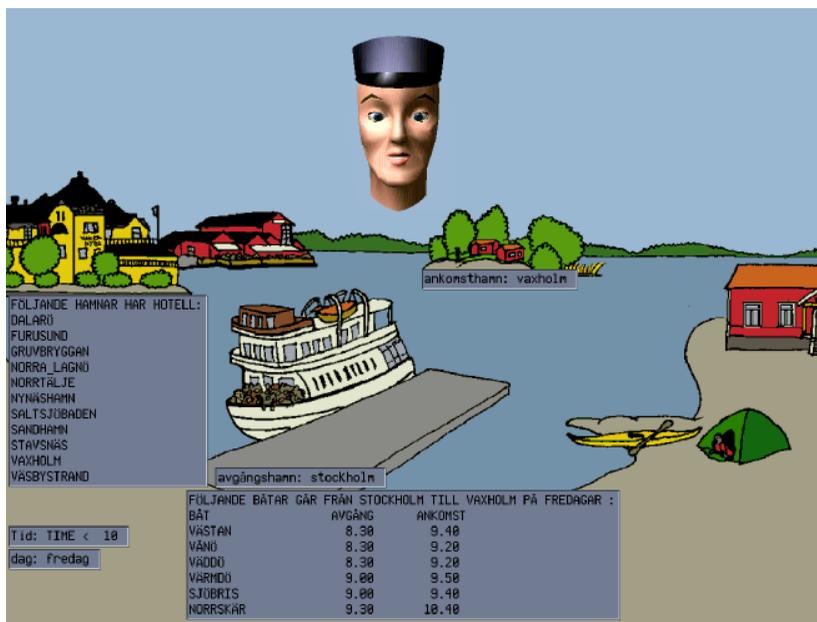


Figura 56: exemplo de agente embutido em sistema de dialogo falado
Fonte: Beskow (2003)

Jemni, Elghoul (2008) descrevem um sistema para produção de sinais utilizando uma abordagem colaborativa. O sistema utiliza o WebSign (JEMNI; ELGHOUL, 2007; JEMNI; ELGHOUL; MAKHLOUF 2007b; CHABEB; ELGHOUL; JEMNI, 2006), onde a

entrada do sistema é um texto e a saída é a interpretação deste texto por um avatar, utilizando um dicionário de sinais.

Relativo à síntese de línguas de sinais com agentes virtuais, Karpouzis et al. (2007) apresentam um sistema que

foca na habilidade de gerar frases de sinais que respeitem as regras da Língua de Sinais Germânica em um grau de precisão que permita seu reconhecimento por sinalizantes nativos como declarações corretas da língua (KARPOUZIS et al. 2007, p. 57).

No contexto da acessibilidade na televisão digital, Souza, Machado e Tavares (2010) afirmam que

Um recurso importante na norma brasileira de TVDI é a transmissão e apresentação de conteúdo com legendas em Libras. A transmissão de conteúdo nesta língua oferece acessibilidade dos programas de TV aos surdos. A apresentação deste conteúdo ao espectador ainda não foi definida. Um dos candidatos a resolver este problema é o uso de avatares 3D para interpretação dos sinais enviados pelas emissoras. Deste modo a incorporação de sistemas 3D ao middleware pode apoiar sistemas para a apresentação de conteúdo em Libras (SOUZA; MACHADO; TAVARES, 2010, p.144, tradução nossa).

O princípio aplicado à síntese de sinais por meio de avatares é análogo ao de reconhecimento e síntese de fala (KURODA et al., 2004), conforme ilustrado na figura 57. Nota-se também a possibilidade de reconhecimento de voz e respectiva tradução para geração de LS, como apresentado na sessão anterior.

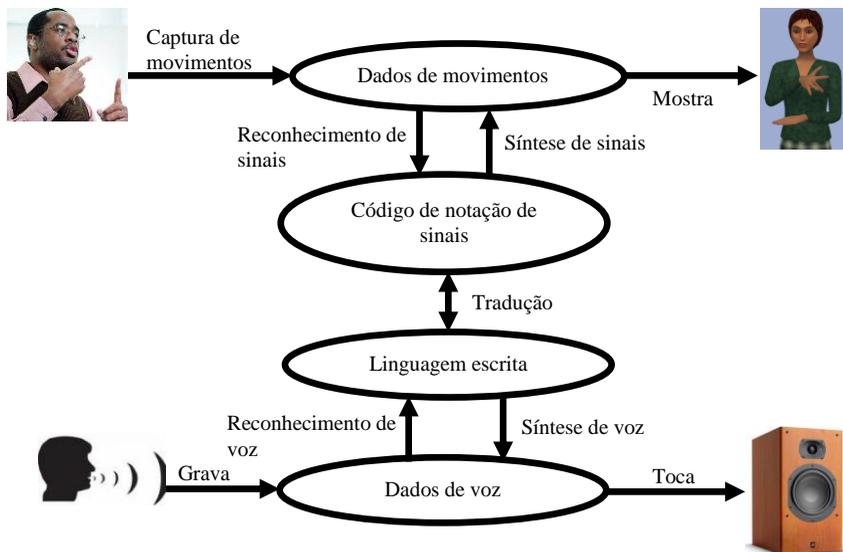


Figura 57: Visão geral de sistemas de sinais que utilizam avatares
 Fonte: Traduzido de Kuroda et al. (2004)

Essencialmente quando se fala em avatares apresentando LS, retrata-se um cenário onde um sinalizante virtual apresenta o resultado de um procedimento de conversão de texto em gestos dos sinais. Em relação à síntese de sinais, Havasi e Szabó (2005) afirmam que

A síntese de sinais pode ser agrupada em dois tipos: **articulatório**, onde a saída é sintetizada por regras que objetivam corresponder ao processo pelo qual humanos articulam a fala; e **concatenativa**, onde sinais pré-gravados (movimentos) são repartidos em pequenos segmentos que podem ser recombinados para criar novos sinais (HAVASI; SZABÓ, 2005, p.637).

Os meios de apresentação das LS revelam diferentes níveis de abstração para apresentação de informações visuais (ADAMO-VILLANI; BENI, 2005). Um meio de representação é o semantróide, que é um avatar reduzido de modo a maximizar o conteúdo semântico enquanto permite seu redimensionamento na tela. Um exemplo de avatar e de semantróide são apresentados na figura 58.



Figura 58: Exemplo de um avatar e um semantróide.

Fonte: Adamo-Villani e Beni (2005, p.3).

O nível de abstração, e conseqüente esforço para interpretação, é maior quando se utiliza a escrita de sinais e menor quando utilizado um humano ou avatar para a apresentação de mensagem em LS (ADAMO-VILLANI; BENI, 2005). O quadro 8 demonstra estas relações.

Quadro 8: Comparação entre diferentes artefatos para comunicação em LS

Fonte: Traduzido de Adamo-Villani, Beni (2005)

Meio de apresentação	Nível de abstração	Movimento	Tamanho
Humano ou avatar	Baixo (concreto)	Sim	Grande
Semantróide	Médio	Sim	Pequeno
Escrita de sinais	Muito alto (simbólico)	Não	Pequeno

Um modelo para o desenvolvimento de legendas em línguas de sinais para TV digital utilizando avatares é apresentado por Brito e Pereira (2009). O modelo sintetiza as etapas de reconhecimento da língua de sinais em um sistema de escrita, o qual permite a edição de conteúdo e a geração da animação dos sinais e sua transmissão no ambiente de TV digital. E baseia-se no uso de técnicas articulatórias, como definido por Maldonado (2007), prevendo a entrega de conteúdo em Libras em duas formas: por envio de vídeo pré-renderizado ou pelo

envio de dados de animação que servem para a geração da animação dos gestos da língua de sinais no próprio receptor do usuário.

A síntese de sinais por meio de avatares precisa considerar os meios que serão utilizados para armazenamento do material produzido, a forma com que os avatares devem ser modelados e elementos que favoreçam o realismo e inteligibilidade da mensagem apresentada.

O próximo tópico trata de padrões e especificações de dados que podem ser utilizados para a animação de avatares.

Posteriormente o uso de técnicas de animação articulatórias e concatenativas são vistas em mais detalhes.

3.5.3.1 Padrões para dados de animação de avatares

Diferentes padrões técnicos existem para especificar a movimentação de um humano virtual, entre as mais importantes estão a especificação H-Anim, a BVH e a VHML.

Gustavsson, Stringlund, Wiknertz (2002) apresentam a verificação, validação e avaliação da VHML (Virtual Human Markup Language). A VHML controla ‘humanos virtuais’ em relação à fala, animação facial, gestos faciais e animação de corpo. A VHML subdivide-se nas linguagens *Emotional Markup language* (EML), *Gesture Markup Language* (GML), *Speech Markup Language* (SML), *Facial Animation Markup Language* (FAML)(VHML, 2001). Nota-se que para a apresentação de sinais em LS é necessário estender a *Body Animation Markup Language* (BAML), pois esta não oferece todos os elementos necessário à representação das LS.

O formato BVH (Biovision Hierarchical Data) é um padrão de dados inicialmente desenvolvido para representar os dados de movimentos capturados a partir de luvas de dados. Atualmente é utilizado no armazenamento de dados de animação de personagens.

A especificação H-Anim (H-Anim, 2011) oferece um conjunto de definições e tipos de dados para o controle de pontos no corpo e na face de humanoides virtuais modelados em VRML e X3D. Os pontos representados pela versão 1.1 da especificação podem ser vistos na figura 59.

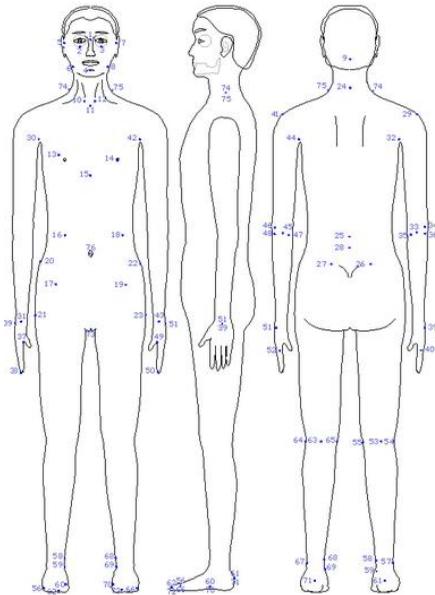


Figura 59: pontos de articulação da especificação H-anim.
Fonte: H-Anim (2011)

Outro padrão de dados utilizado na codificação de dados de movimentos é a MPEG4 *Body Animation Parameters* (BAPs) e *Face Animation Parameters* (FAPs) (PREDA; PRETEUX, 2002).

Com estes padrões diferentes juntas de um esqueleto padrão podem ser manipuladas por meio de informações enviadas em um fluxo de dados com os BAPs e FAPs. Entretanto algumas expressões faciais não podem ser implementadas na animação de LS pois não existem FAPs para elas (PAPADOGIORGAKI et al. , 2004).

A síntese de sinais com técnicas de animação concatenativa realiza a junção de unidades de conteúdo em LS representadas com os formatos de dados apresentados nesta sessão. Uma visão sobre este tipo de síntese de sinais é apresentada na próxima sessão.

3.5.3.2 Síntese concatenativa

A síntese concatenativa consiste na junção dos dados de animação armazenados em uma base de dados sobre sinais. Estes dados utilizam os padrões descritos no tópico anterior e possibilitam o reuso de dados provenientes da captura de movimentos de um intérprete ou então de dados gerados manualmente.

Um sistema para autoria de conteúdo em LS baseado em avatares tridimensionais pode ser composto por três componentes principais: um componente de apoio à importação de modelos 3D; um componente de apoio à animação e outro para apoio à renderização (ADAMO-VILLANI et al., 2010).

O componente de apoio à animação

permite ao usuário importar sinais de uma base de sinais, criar novos sinais, criar articulações faciais, realizar a ligação dos sinais e articulações faciais de modo a gerar um discurso contínuo, e digitar um script em ASL que gera uma animação correspondente. (ADAMO-VILLANI et al., 2010, p. 39)

Os sinais cadastrados são associados à palavras da língua oral, de modo que um texto de entrada é associado a um dicionário de sinais ou glossário, e caso haja um registro no dicionário correspondente ao fragmento de texto, a animação do sinal correspondente é apresentada. Caso o sinal não esteja cadastrado no dicionário, o texto é fragmentado e apresentado pelo agente 3D utilizando a datilologia (PHAN; NGUYEN; BUI, 2009).

O avatar apresentando sinais gerados por meio da técnica por Phan, Nguyen e Buy (2009) é apresentado na figura 60.

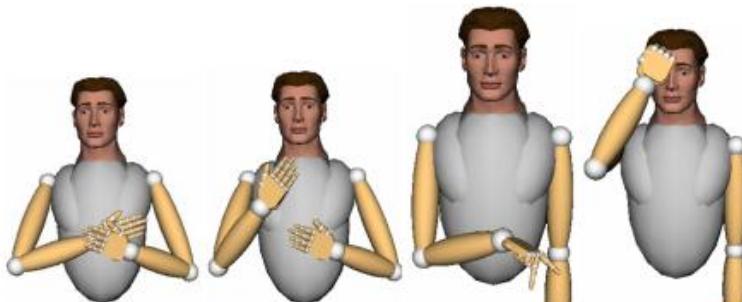


Figura 60: agente 3D interpreta a frase “Eu vou para a escola de manhã” em LS vietnamita.

Fonte: Phan, Nguyen e Bui (2009, p. 326)

Para tornar mais prática a tarefa de edição da animação durante o cadastro dos sinais no glossário, Jemni e Elghoul (2008) propõem a animação direta do avatar, sem o uso de um sistema de notação intermediário. Nesta abordagem utilizam a codificação do sinal numa linguagem de definição de sinais orientada ao avatar, a SML (*Sign Markup Language*). A partir da SML é gerado um fluxo de dados de animação baseados na especificação H-Anim que servirá para movimentar o avatar de acordo com os sinais que se deseja realizar. Um exemplo de aplicação pode ser visto na figura 61.

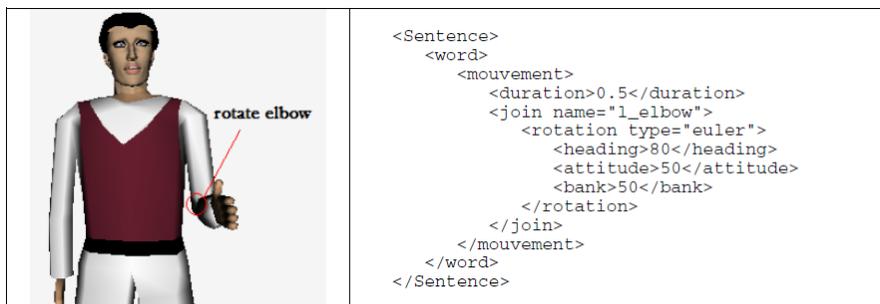


Figura 61: exemplo de SML(X)

Fonte: Jemni, Elghoul (2008, p. 675)

A animação direta é feita por meio de um aplicativo web, onde a dimensão de tempo (duração) é adicionada na representação de um sinal. A interface de cadastro de sinais de Jemni e Elghoul (2008) pode ser vista na figura 62. Nesse sistema,

O procedimento de criação de um sinal é resumido em três etapas: primeiramente se escolhe a posição do quadro inicial na linha do tempo arrastando-se o ponteiro para a posição selecionada. Em seguida deve ser escolhido a junta que deve ser rotacionada utilizando-se o seletor de juntas. E finalmente o usuário deve aplicar as barras de rotação para indicar o valor da rotação da junta, repetindo o processo até a

criação do sinal (JEMNI; ELGHOUL, 2008, p.675, tradução nossa).

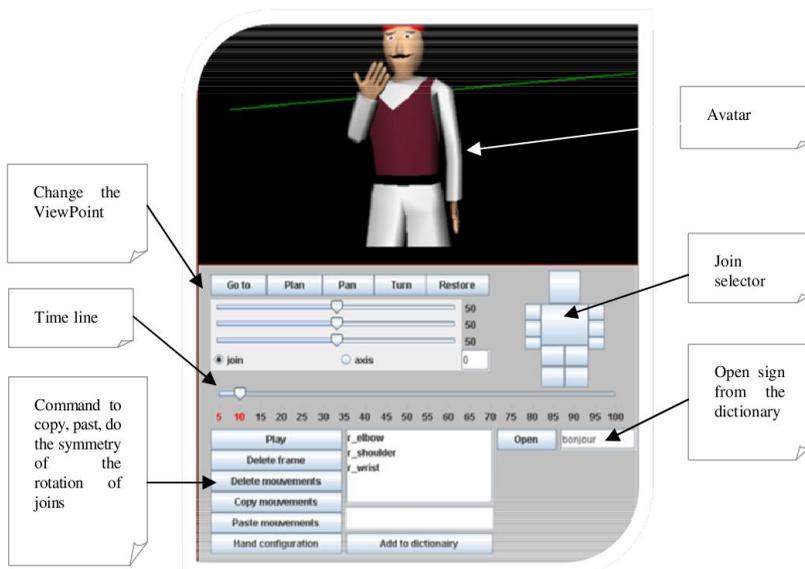


Figura 62: Interface para animação direta dos sinais em um avatar
Fonte: Jemni, Elghoul (2008, p. 676)

O vídeo de um intérprete realizando os sinais pode ser usado como uma referência para a criação das animações do avatar (ADAMO-VILLANI; BENIL, 2005).

A síntese de sinais é utilizada em um sistema de produção de conteúdo educativo para surdos, onde sinais são gerados automaticamente a partir de descrições textuais e interpretados num navegador web por meio de avatares (GHOUL; JEMNI, 2009).

A arquitetura baseada no método de Jemni, Elghoul (2010) é aplicado em um site de notícias sobre esportes, que gera conteúdo em LS sob demanda a partir do texto em língua oral. A arquitetura é apresentada na Figura 63.

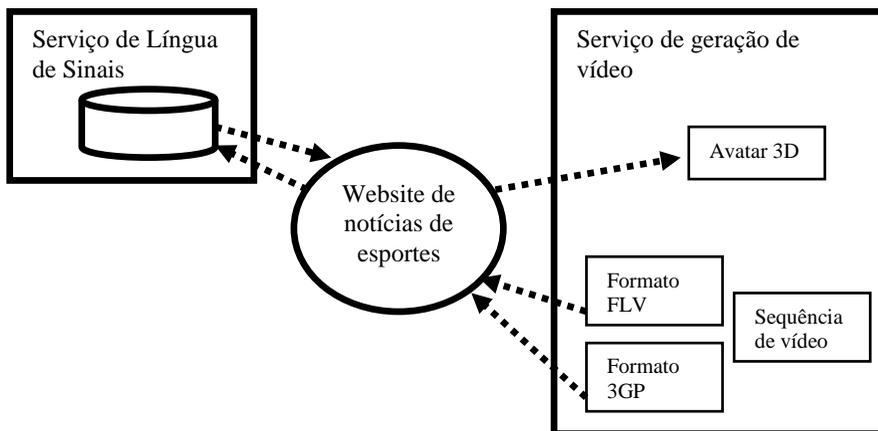


Figura 63: arquitetura de um serviço de geração de vídeos com avatares sinalizantes

Fonte: Adaptado de Othman, Ghoul, Jemmi (2010, p. 164).

De modo similar, Xu et al. (2008) apresentam um sistema de animação de avatares para a LCS (Língua Chinesa de Sinais) a partir de um editor de animações e uma base de sinais com seus dados de movimentos. O sistema é descrito na figura 64, onde um cliente fim (*Terminal Client*) requisita uma sentença em LS a um aplicativo (*3rd party application*) que converte a requisição em uma representação XML enviada a um serviço web de língua de sinais (*CSL WebService*).

A sentença enviada ao CSL WebService gera uma representação XML segmentada em palavras (*Word Segmented Module*) que são enviadas para animação por um módulo de síntese de LS (*CSL Synthesis Module*). Este módulo gera dados de animação de avatar que são retornados ao cliente fim.

Para produzir os dados de animação, o módulo de síntese de LS consulta uma base de dados (*CSL Database*) que foi previamente alimentado com dados de animação gerados por uma ferramenta de síntese de animações (*Animation Synthesis Tool*) a partir de um editor de posturas corporais e configuração de mão (*Hand Shape and Static Posture Editor*).

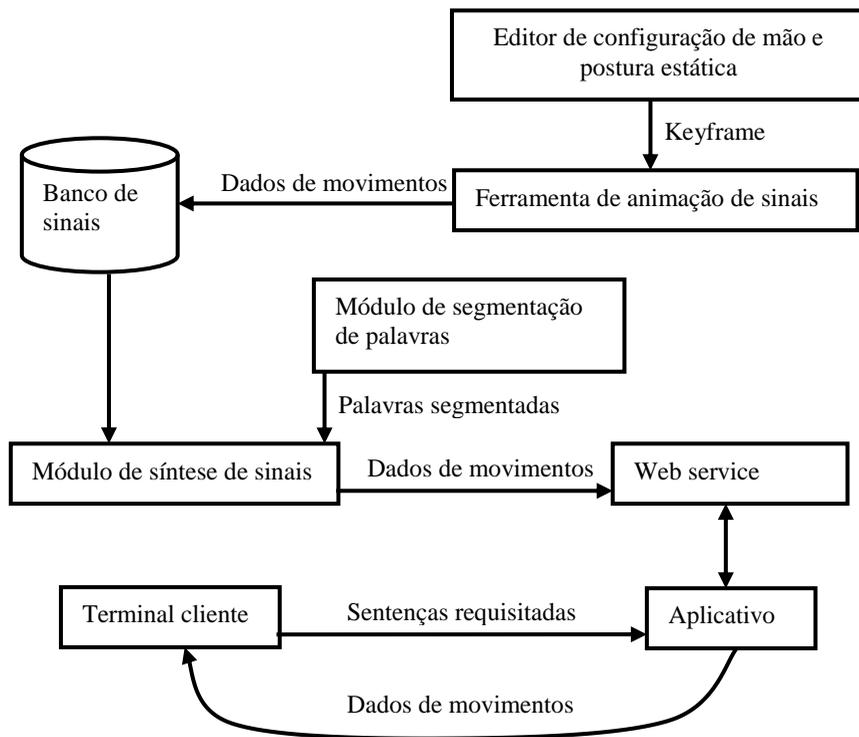


Figura 64: Arquitetura de Web Service para síntese de sinais em CSL (Chinese Sign Language)

Fonte: Xu et al. (2008 , p.448).

Kaneko et al. (2010) usam um script do TVML (Television Markup Language) para animar um avatar em língua de sinais. A animação é gerada pela junção de arquivos.bvh (*Bio Vision Hierarchy*) gerados a partir de técnicas de captura de movimento (MITOBE, 2006; HAVASI; SZABÓ, 2005).

Além dos arquivos .bvh , é mantida uma base de metadados que indica quais as partes de cada arquivo .bvh contém os movimentos essenciais para a compreensão do significado do sinal de forma que transições entre sinais podem ser manipuladas para a junção de diferentes arquivos na geração de uma sequência contínua. O sistema

entretanto não permite a animação de expressões faciais, limitando a comunicação em língua de sinais.

A animação concatenativa por *keyframes* não permite apresentar adequadamente as transições entre os sinais, e dificulta a inclusão de sinais não previstos ao conteúdo, conforme destacam Januário, Leite e Koga (2010).

[..] quando um personagem precisa realizar um movimento localizado, como abrir e fechar sua mão, este é um movimento pré-definido e usado com frequência, logo a solução nestes casos é, em tempo de modelagem, gerar animações dos movimentos da mão se abrindo ou fechando e carregar essa animações em tempo de execução; no [..] caso [das LS], as possibilidades de transição, de uma configuração de mão para outra, são tantas que tal técnica se torna inviável, e por isso é preciso a modificação do modelo [3D do avatar] em tempo de execução, seja pela manipulação dos ossos (*bones*), seja pela manipulação da própria malha (*mesh*) (JANUÁRIO; LEITE; KOGA, 2010, p. 64).

Diante a necessidade de criação de novas transições entre sinais em tempo de execução, as técnicas associadas à síntese articulatória demonstram-se mais adequadas, as quais são descritas no próximo tópico.

3.5.3.3 Síntese Articulatoria

A síntese articulatória utiliza os componentes fonológicos dos sinais representados num sistema de escrita. Tanto os sistemas de escrita da língua oral, quanto em LS, estes dois sistemas podem ser utilizados como parâmetros para a animação de avatares. Entretanto o uso de um sistema de escrita oral para gerar animações em um sistema visual-espacial possui limitações e estão mais associados à síntese concatenativa.

Duas abordagens para a geração de sinais a partir de notação escrita são apresentadas nas figuras a seguir: o uso do HamNoSys e sua codificação em SigML (*Sign Gesture Markup Language*), na figura 65, e o uso do *SignWriting* e sua codificação em SWML (*Sign Writing Markup Language*)(COSTA; DIMURO, 2002), conforme figura 66 (JEMNI; ELGHOUL, 2008).

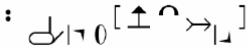
	<pre> <sigml> <hns_sign gloss="DGS_going-to"> <hamnosys_manual> <hamsympar/><hamfinger2/> <hamthumboutmod/><hamextfingeruo/> <hampalml/><hamparbegin/> <hammoveo/><hamarcu/> <hamreplace/><hamextfingerdo/><hamparend/> </hamnosys_manual> </hns_sign> </sigml> </pre>
---	--

Figura 65: exemplo de HamNoSys e SigML

Fonte:Jemni, Elghoul (2008 ,p. 67)

	<pre> <signbox> <symp x="46" y="37" x-flop="0" y-flop="0" color="0,0,0"> <category>04</category> <group>02</group> <symbnum>001</symbnum> <variation>01</variation> <fill>01</fill> <rotation>04</rotation> </symp> </signbox> </pre>
---	---

Figura 66: exemplo de SignWriting e SWML

Fonte:Jemni, Elghoul (2008 , p. 673)

Papadogiorgaki et al. (2004) e Papadogiorgaki, Grammalidis e Makris (2005) descrevem a aplicação do *SignWriting* na animação de avatares onde utilizam a especificação H-Anim, no projeto *VSigns*. De acordo com Papadogiorgaki et al., (2004), o aplicativo desenvolvido pode ser aplicado

tanto como com “plug-in” de aplicativos existentes (i.e. dicionários de línguas de sinais) ou como uma ferramenta independente para a criação de animações para notícias de TV (i.e. previsões do tempo) (PAPADOGIORGAKI et al., 2004, p.266, tradução nossa).

O fluxograma do algoritmo de geração de sinais por meio de avatares a partir de *SignWriting* é descrito na figura 67.

Cada sinal está contido em uma *SignBox*, onde cada símbolo possui:

- a) Um número inteiro especificando a forma do símbolo;
- b) Um parâmetro de variação (0 ou 1 para símbolos da mão; 1, 2 ou 3 para símbolos de movimentos e pontuação) que especifica possíveis variações (transformações complementares) do símbolo;
- c) Um parâmetro de preenchimento (0, 1, 2, ou 3 para símbolos da mão ou de pontuação; 0, 1, 2 para símbolos de movimento) que especificam como a forma é preenchida, indicando como ela é vista pelo sinalizante;
- d) um parâmetro de rotação (de 0 a 7) que especifica uma rotação no sentido anti-horário aplicada ao símbolo, em passos de 45 graus;
- e) um parâmetro de inversão (0 ou 1) que indica se o símbolo é verticalmente espelhado ou não, em relação ao símbolo básico, e finalmente;
- f) As coordenadas x e y do símbolo dentro da *SignBox* (PAPADOGIORGAKI; GRAMMALIDIS; MAKRIS, 2005, p.152).

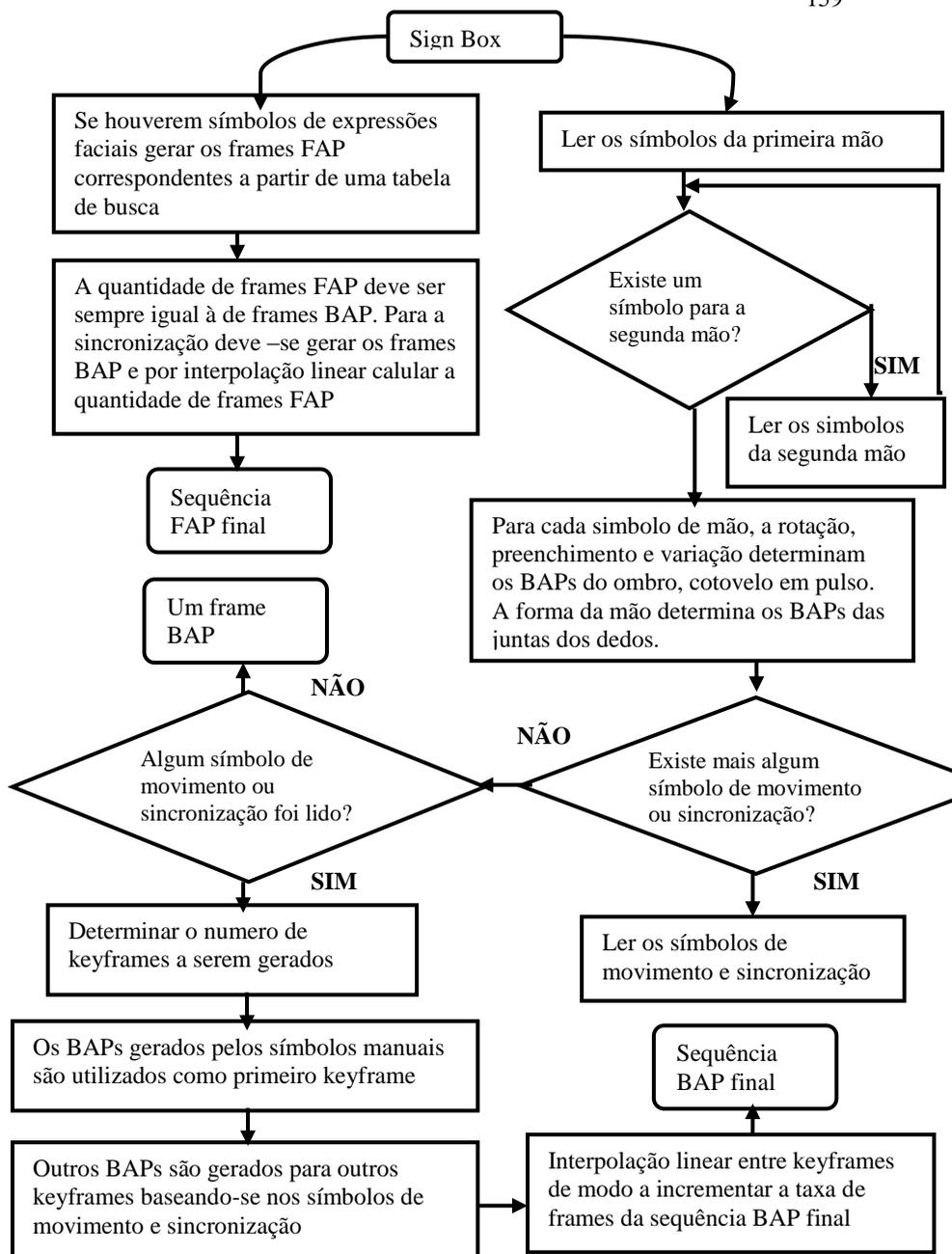


Figura 67: sequência para animação de sinais com H-Anim a partir de SignWriting

Fonte: Papadogiorgaki et al. (2004)

Este animador de gestos é utilizado no aplicativo *VSigns*, que

gera sequências de animação VRML da notação de línguas de sinais, baseada na Animação de Corpo MPEG-4, onde a notação de sinais, em *SignWriting*, é fornecida como entrada e inicialmente convertida em SWML. Cada sinal básico do *SignBox* é convertido em uma sequência de Parâmetros de Animação de Corpo (BAPs) do padrão MPEG-4, correspondente ao gesto representado. Por fim, estas sequências são usadas para animar os avatares VRML em conformidade com H-ANIM, reproduzindo os gestos exatos representados na notação da língua de sinais. Por ser um sistema de animação vetorial que usa poses-chave, o animador demanda recursos computacionais para calcular interpolações a fim de gerar as poses intermediárias (DENARDI; MENEZES; COSTA, 2004, p.4).

Para a geração da animação,

Todos os símbolos manuais, faciais e de movimento que foram informados na entrada no formato SWML são convertidos em parâmetros de animação. Para produzir uma animação, os parâmetros de animação são selecionados seguindo a ordem de leitura dos símbolos em um sinal, como determinado pela sequência de soletração do sinal [...]. A animação é apresentada como uma sequência de poses em quadros-chaves, expressões faciais e formas manuais. A animação foi conseguida pela junção destas poses. Isto é feito usando interpolação de quartênios de modo a obter-se uma transição suave entre os quadros. O avatar é então exibido na tela realizando a animação dos sinais correspondentes à entrada (MOEMEDI, CONNAN, 2010, p.5).

Um exemplo de animação facial gerada a partir de *SignWriting* é apresentado na figura 68.

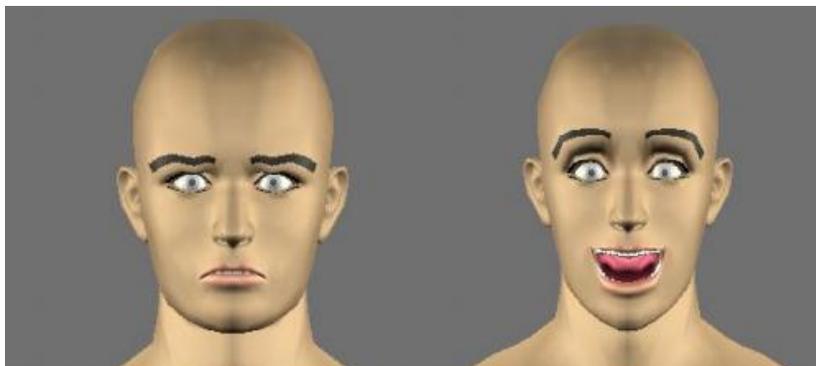


Figura 68: exemplo de animação de expressão facial a partir de escrita de sinais.
Fonte: Moemedi, Connan (2010, p.4).

Outras técnicas para armazenamento e apresentação de língua de sinais permitem que o autor do conteúdo escreva a sequência de animação utilizando o sistema *HamNoSys* (ZWITSERLOOD et al. 2005; SAN-SEGUNDO et al.;2008).

A síntese articulatória de LS pode ser integrada em materiais multimídia para aprendizagem como demonstram Verlinden, Zwitserlood e Frowein (2005) ao aplicarem o sistema de notação de língua de sinais *HamNoSys* (*Hamburg Notation System*) codificado na *Signing Gesture Markup Language* (SiGML), linguagem que funciona como uma descrição da geometria que o avatar deve assumir para cada sinal.

Outro sistema voltado à animação articulatória das LS é o Visicast (2011), que é um projeto voltado à produção e apresentação de LS por meio de avatares. Os fundamentos aplicados neste projeto são apresentados por Elliot et al.(2007) e Kennaway et al.(2007).

Elliot et al. (2007) descrevem a modelagem linguística e as tecnologias para processamento de linguagem necessárias para a apresentação de língua de sinais por meio de avatares, e enumeram as seguinte etapas para a síntese de sinais a partir do sistema de notação/escrita *HamNoSys*:

- Implementação das formas manuais que podem ser descritas pelo *HamNoSys* em um modelo de cinemática inversa;

- Estabelecimento de regras de tradução das especificações de posicionamento das mãos em relação às coordenadas numéricas do espaço tridimensional de sinalização;
- Identificação das configurações naturais e movimentos de partes da anatomia do sinalizante;
- Geração de movimentos naturais, com acelerações e desacelerações, e evitar colisões ou sobreposições entre as partes do corpo.

Para tratar as limitações do *HamNoSys* quanto à notação de expressões faciais, Elliot et. al (2007) estabelecem um padrão adicional para definir as características não manuais. Os resultados são aplicados nos projetos Visicast e eSign por meio do avatar VGuido, apresentado na figura 69.



Figura 69: Avatar VGuido
Fonte: eSign

A geração de uma animação a partir de HamNoSys segue as seguintes diretrizes:

- As categorias discretas de significados tais como grande, pequeno, rápido e devagar, localização de pontos com nomes HamNoSys, devem ser traduzidas em quantidades numéricas apropriadas a cada avatar;

- Utilização de cálculos de cinemática inversa ou rotações de juntas de braços apropriadas, movimentos de torso devem ser determinados de modo a colocar as mãos onde requeridas;
- Todas outras informações que não podem ser registradas por meio da HamNoSys devem ser inferidas. Sinais registrados por leitores humanos geralmente omitem informação que é óbvia para um leitor humano mas difícil de calcular automaticamente. Por exemplo o contato entre mãos que é geralmente definida por um símbolo específico do HamNoSys sem especificar os pontos específicos nas mãos que são colocados em contato;
- Colisões e penetrações entre partes do corpo devem ser evitadas;
- Descrições de movimentos faciais devem ser traduzidos em deformações temporais da superfície da malha do avatar;
- As posturas e movimentos devem ser precisos o suficiente para serem legíveis como sinais, e naturais o suficiente para serem aceitas pelos usuários finais;
- Toda a computação [da animação] deve ocorrer em tempo real, utilizando uma proporção pequena do tempo suficiente para permitir a renderização do avatar (KENNAWAY et al., 2007, p.33, tradução nossa) .

Como vantagem do sistema Visicast (2011), Kennaway, Glauert e Zwitserlood (2007) afirmam que:

Tanto a definição dos sinais na nossa linguagem de notação e a composição das sequências de sinalização podem ser realizadas por uma pessoa em um computador. Não é necessário equipamento de captura de movimentos ou de vídeo (KENNAWAY; GLAUERT; ZWITSERLOOD, 2007).

Além da simplicidade dos equipamentos utilizados em termos de hardware, o uso combinado de sistemas de notação de sinais e avatares oferece algumas vantagens sobre o vídeo tradicional (KENNAWAY; GLAUERT; ZWITSERLOOD, 2007):

- A continuidade não é um problema, pois qualquer parte da sinalização pode ser animada por qualquer avatar;
- Detalhes do conteúdo podem ser editados sem a necessidade de regravação de sequências;
- São mínimos os requisitos de dados para transmissão e armazenamento, pois, o software na máquina do cliente traduz a notação de script em dados de animação e renderiza o vídeo do avatar, deste modo apenas a notação em script precisa ser transmitida e pode ser muito menor do que dados de vídeo comprimido ou dados de movimentos (KENNAWAY et al. 2007, p. 17, tradução nossa).

Os dados de animação podem ser transmitidos a partir do servidor são independentes da taxa de frames por segundo do vídeo final na máquina do cliente, de modo que esta pode ser ajustada de acordo com sua capacidade de renderização. Também, ao se transmitir a informação sinalizada na forma de uma notação que serve para gerar uma animação no receptor do usuário a quantidade de dados necessária para a transmissão é muito inferior à requerida para dados de vídeo, mesmo que altamente comprimidos (KENNAWAY; GLAUERT; ZWITSERLOOD, 2007).

Além disso, o usuário pode ter controles adicionais que não são possíveis com o vídeo tradicional, por exemplo o ângulo de visualização pode ser ajustado durante a exibição, ou então é possível a escolha de diferentes avatares (KENNAWAY et al. 2007).

Se o conteúdo é originalmente formulado como texto escrito, torna-se necessário determinar um meio apropriado de apresentar esta informação por meio de sinalização. A sinalização escolhida pode ser representado por uma sequência de glosas, que são palavras usadas como etiquetas para os sinais correspondentes (algumas vezes

parametrizadas com informações que podem variar entre um uso do sinal ou outro). Cada glosa é associada com a notação SiGML que pode ser usada para sintetizar a animação necessária (KENNAWAY et al. 2007, p. 19).

O *SignSynth* (GRIEVE-SMITH, 2001) é um sintetizador de sinais que considera um texto de entrada em notação de línguas de sinais (Stokoe) e o traduz em uma estrutura de representação interna, que é então convertida na animação de um avatar que utiliza VRML e arquivos Web3D. O avatar é apresentado na figura 70 a seguir.



Figura 70: avatar do SignSynth
Fonte: SignSynth (2011)

Estas abordagens de tradução automática buscam resolver o problema de gravação do intérprete e diminuem a quantidade de dados para armazenamento ou transferência das LS.

O tópico seguinte analisa a aplicação de avatares bidimensionais na comunicação por sinais.

3.5.3.4 Avatares bidimensionais

Avatares bidimensionais simplificados podem ser utilizados na apresentação de LS (SAN-SEGUNDO, 2008). A figura 71 apresenta um exemplo de avatar bidimensional. Em relação a seu uso, San-Segundo afirma que

Embora não tenhamos realizado uma avaliação completa de como surdos percebem os gestos de nossos agentes, avaliações preliminares com letras do alfabeto revelam que pessoas surdas encontram dificuldade de entendimento em menos de 30% dos gestos. Essa situação ocorre apenas na primeira vez que o gesto é realizado. Nas vezes subsequentes, dado que o gesto é realizado sempre da mesma forma, os surdos reconhecem-no facilmente (SAN-SEGUNDO et al., 2008, p.537, tradução nossa).

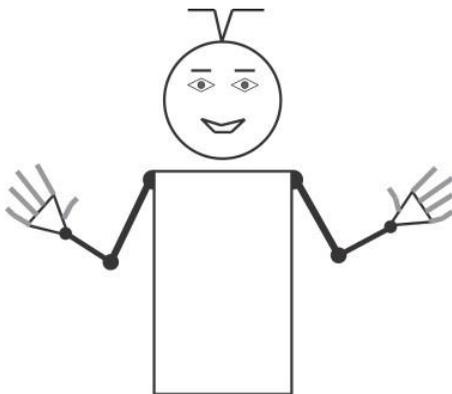


Figura 71: Agente de representação de gestos.

Fonte: San-Segundo et al. (2008).

Um software voltado à acessibilidade no ambiente de TV digital, denominado RybenáTV, extrai as legendas ocultas do conteúdo apresentado realiza uma tradução destas legendas para a Libras e a apresenta por meio de um avatar bidimensional (AMORIN et al., 2010). O aplicativo é apresentado na figura 72.



Figura 72: Software Rybená apresentando legenda em LS
 Fonte: Amorin et al. (2010, p.246)

O sistema é fortemente orientado pelas estruturas da língua oral, de modo que no Rybená

a tradução ocorre na sequência de português sinalizado, obedecendo a uma estrutura linguística diferenciada da LIBRAS em sua gramática própria. Diante de palavras que não constam em seu banco de dados, o sistema utiliza-se do alfabeto manual na transmissão do conteúdo, ou seja, soletra a palavra letra por letra. Desta forma, o funcionamento do aplicativo pode ocasionar problemas de usabilidade no acesso ao conteúdo disponível. Sabe-se que a conversão do português escrito para a LIBRAS, por meio de sistema computacional automático é um grande desafio que ainda precisa ser aprimorado (CORRADI, 2007, p.98)

Outra forma similar de apresentação da língua de sinais é a escrita de sinais em sua forma animada, baseada nas referências escritas dos movimentos executados nos sinais (DENARDI, 2006). Com esta finalidade, Denardi, Menezes, Rocha (2004) apresentam o AGA-Sign, um animador de gestos aplicado a língua de sinais que utiliza a escrita

de sinais *SignWriting* para a representação de sinais. Denardi (2006) descreve o sistema para animação de gestos dos sinais, que utiliza autômatos finitos. A visão geral do sistema é apresentado na figura 73 abaixo.

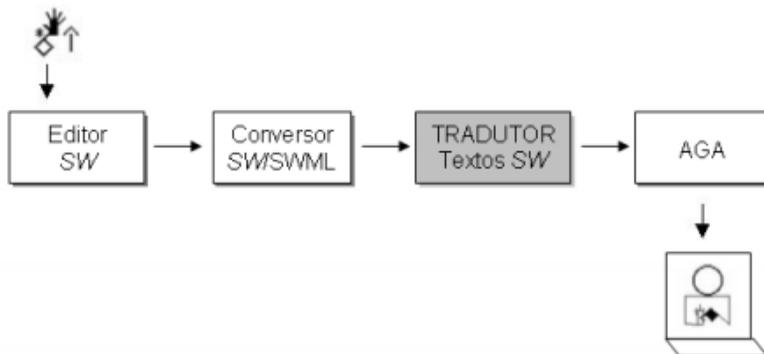


Figura 73: visão geral do sistema de animação de gestos de sinais AGA-Sign
Fonte: Denardi (2006, p.17)

No sistema apresentado, um ‘Editor SW’ permite a entrada de escrita de sinais, que então é representada em SWML. Em seguida, o ‘Tradutor de Textos SW’ converte o SWML em AgaML, uma linguagem que especifica animações apresentadas pelo animador de gestos baseado em autômatos (AGA). Um exemplo de resultado pode ser visto na figura 74, onde três quadros de uma animação são apresentados.

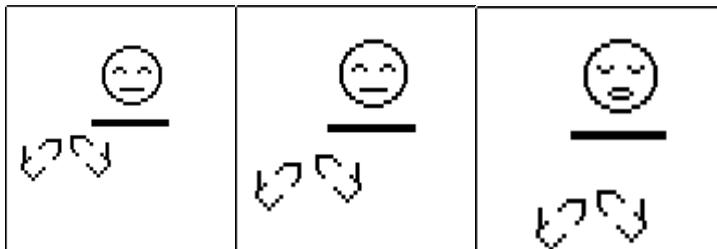


Figura 74: seqüência de animação de sinais a partir de SW
Fonte: Denardi (2006)

Os sistemas baseados em avatares bidimensionais oferecem uma alternativa a mais para a apresentação das LS, e estão associados a um baixo custo computacional.

O próximo tópico descreve um software para tradução que utiliza agentes virtuais.

3.5.3.5 O software de tradução PoliLibras

Este tópico apresenta tradutor português para Libras denominado PoliLibras (JANUÁRIO; LEITE; KOGA, 2010). O tradutor utiliza a síntese articulatória para gerar a animação dos sinais cadastrados em um dicionário segundo os parâmetros fonológicos do *SignWriting*. Já a apresentação das sentenças traduzidas faz uso de síntese concatenativa, unindo as animações dos sinais individuais armazenados no dicionário.

O processo de tradução e síntese gráfica ocorre a partir do envio de um texto em português a um tradutor que requisita a um dicionário os sinais que serão enviados para apresentação por um sintetizador gráfico.

Segundo os autores da plataforma, no caso do PoliLibras,

A tradução é dirigida por sintaxe, o que quer dizer que é realizada uma análise sintática da estrutura frasal portuguesa e então aplicada uma transformação que faz a adaptação da sintaxe do português para a Libras (POLILIBRAS, 2011, sem paginação).

Na tarefa de tradução são utilizados um analisador morfológico e outro sintático, além de um transformador sintático e um contextualizador. Essencialmente,

O analisador morfológico identifica as classes gramaticais das palavras do texto, o que permite ao analisador sintático montar a árvore sintática de cada sentença do texto; uma vez de posse da árvore, o transformador sintático aplica as transformações sintáticas na sentença que fazem a adaptação da estrutura da frase de uma língua para a outra; é nesse momento que a consulta é feita ao dicionário, sendo que depois dessa consulta a sentença ainda passa por uma última, o contextualizador, que vai realizar mais algumas adaptações, como a conjugação de verbos na Libras, dentre outros (POLILIBRAS, 2011, sem paginação).

A figura 75 a seguir descreve a arquitetura do sistema de tradução PoliLibras.

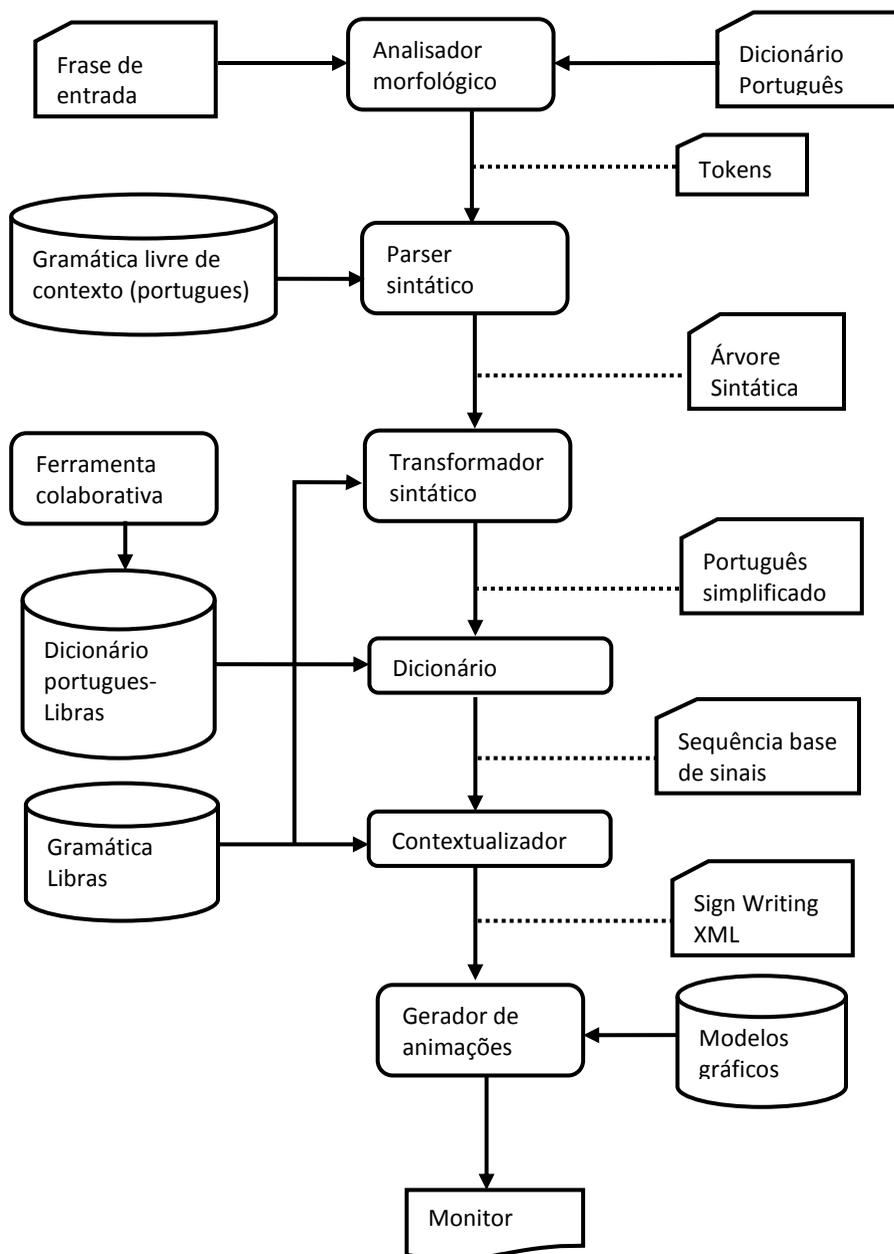


Figura 75: Arquitetura do sistema de tradução PoliLibras
 Fonte: PoliLibras (2011).

O tradutor aceita como entrada frases de período simples em ordem direta, de modo que frases com estruturas diferentes são traduzidas palavra por palavra.

Os termos são traduzidos utilizando-se o dicionário Wiki-Libras (WIKI-LIBRAS, 2012).

A modelagem morfológica dos sinais é baseada no *SignWriting* e utiliza um tradutor de SWML para o XML específico do aplicativo que apresenta os sinais utilizando síntese articulatória.

Para apresentar o conteúdo em LS é utilizado o *Virtual Jonah*, que consiste em um sintetizador de sinais que utiliza computação gráfica.

O avatar realiza os sinais descritos em um arquivo texto, que consiste em um XML que descreve os sinais de acordo com nosso modelo do *SignWriting*. *Virtual Jonah* está disponível como uma biblioteca para ser utilizado em aplicações que precisem de uma interface em LIBRAS (POLILIBRAS, 2011, sem paginação).

De fato, o *Virtual Jonah* consiste num semantróide, o qual utiliza como API gráfica a biblioteca Java Processing e OpenGL, e pode ser visto na figura 76.

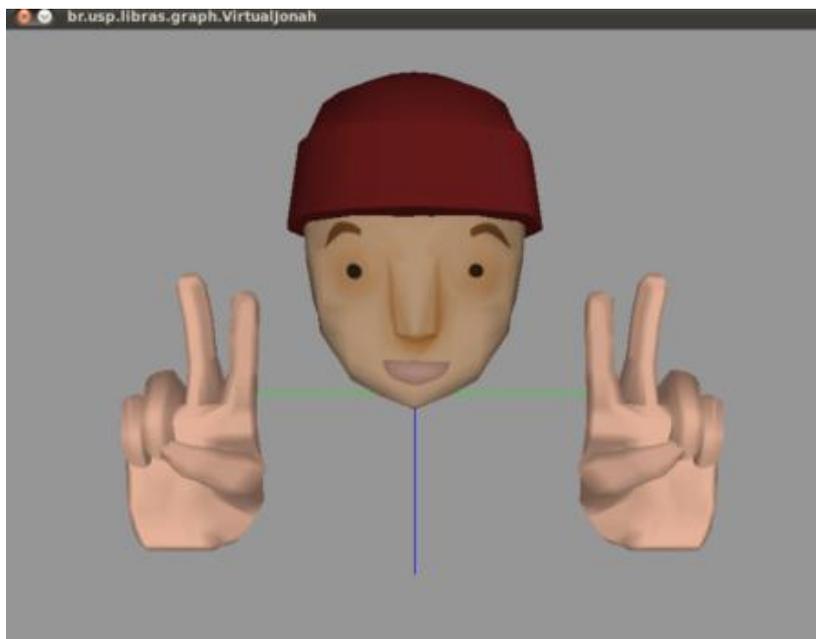


Figura 76: Virtual Jonah apresentando LS
Fonte: <http://code.google.com/p/virtual-jonah/>

O dicionário português/Libras, que é utilizado na tradução, permite a edição colaborativa e a pesquisa por palavras em Libras a partir de palavras em português. Na edição colaborativa os usuários podem cadastrar novos verbetes ou editar os existentes, permitindo a expansão do vocabulário. O dicionário permite também que outras aplicações acessem as informações cadastradas através de uma arquitetura de serviços na web.

3.5.3.7 *Questões relativas à qualidade da síntese de sinais em avatares*

A síntese de sinais a partir de um agente virtual busca gerar movimentos que sejam equivalentes aos de um intérprete humano, entretanto alguns fatores precisam ser considerados para que a sinalização seja realista e cumpra com seu objetivo de comunicação.

Os requisitos que a animação a partir de avatares deve atender incluem as de um intérprete humano, como cor de roupa, iluminação do estúdio e codificação do vídeo. Além dos requisitos de um TILS real, a síntese por meio de avatares deve considerar outros fatores de modo a evitar movimentos robotizados e falta de emoção na entonação dos sinais.

Ao avaliarem a aplicação da técnica de animação de sinais em páginas web, Verlinden, Zwitterlood e Frowein (2005) observaram que a compreensão dos sinais foi prejudicada pela ausência de expressões faciais, e que o avatar, apesar de realístico, apresentou movimentos considerados ‘robóticos’. Além disso, o controle de velocidade da animação demonstrou-se uma funcionalidade importante.

A síntese de língua de sinais por técnicas articulatórias exige a definição dos *keyframes* e sua interpolação para apresentação de movimentos e transições entre sinais, sendo que estas transições podem apresentar comportamentos não naturais e robóticos (MALDONADO, 2007).

Na animação por *keyframes* apenas os quadros principais são definidos na animação e os quadros intermediários são gerados por meio do cálculo das posições intermediárias do objeto em movimento. Na geração dos frames intermediários podem ocorrer comportamentos fisicamente incoerentes, devido a não detecção de colisões entre os sólidos envolvidos (ADAMO-VILLANI et al., 2010). Para evitar estas colisões recomenda-se a adição de *keyframes* intermediários de modo a corrigir o movimento do avatar (WOLF et al., 2006).

Para tornar a animação por *keyframes* de cada sinal mais realista, Adamo-Villani e Beni (2005) descrevem um sistema onde um sinalizante realiza os sinais e gera um registro em vídeo. Este vídeo é importado em um software 3D, onde é usado como referência para a criação das animações.

Em relação à aparência dos avatares, Adamo-Villani et al. (2009) investigaram a aceitação de sinais realizados por dois tipos de avatares: um altamente segmentado e outro não segmentado, e observaram que os sinais realizados pelo avatar com formas mais suaves (não segmentado)

foram melhor aceitos. Os avatares são apresentados na figura 77 e 78 a seguir.

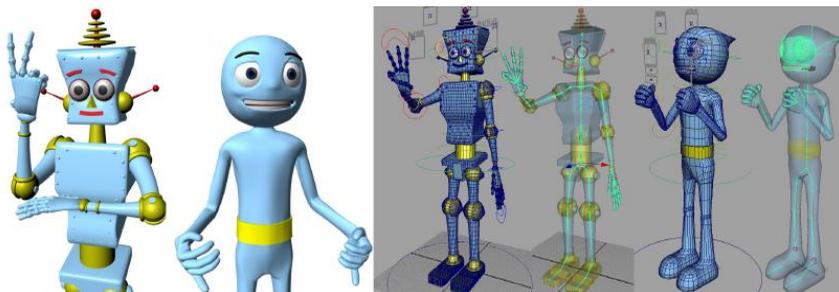


Figura 77: avatares Robby e Torrents renderizados e seus sistemas de esqueletos.

Fonte: Adamo-Villani et al. (2009)



Figura 78: configurações de mão realizadas por um sinalizante humano e por avatar

Fonte: Adamo-Villani et al. (2009)

Aspectos não verbais do avatar são um dos fatores que influenciam a percepção de movimentos dos sinais, aumentando o significado da mensagem. Sua aplicação em avatares sinalizantes mais

convincentes pode aumentar a aceitação pelos usuários, conforme afirma Efthimiou et al. (2009).

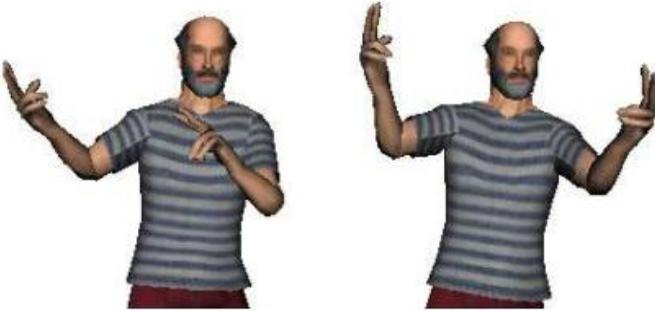
Um dos problemas mais difíceis na síntese de sinais é a conversão da descrição linguística da expressão sinalizada em uma animação suave via cinemática inversa, com o posicionamento adequado das mãos em contato com o corpo, e a geração realística de elementos prosódicos, tais como o stress visual apropriado. Para este fim, o corpus de língua de sinais [...] não deve apenas englobar informações fonéticas e gramaticais, mas também informações prosódicas. [estas informações] Em conjunto com as funcionalidades derivadas do rastreamento visual e do componente de reconhecimento, permitem maior realismo nas animações (EFTHIMIOU et al., 2009, p.25, tradução nossa).

De modo a tratar os aspectos não naturais decorrente da síntese de sinais, Heloir e Gibet (2007) elaboram um modelo teórico sobre a fonologia dos sinais, no qual identificam e representam as variações espaciais, temporais e estruturais encontradas na apresentação de um mesmo sinal. Estas variações são causadas por diferentes estilos utilizados, refletindo a idade, gênero, cultura e estado emocional do sinalizante.

Temporalmente, a variação do estilo do gesto consiste na variação de sua duração entre suas fases: preparação, curso, retração parcial, retração e pausa (*preparation, stroke, partial retraction, retraction, e hold*). Já a variação espacial/estrutural acontece na relação proximal/distal da mão em relação ao corpo, *weak drop* (quando um sinal que usa as duas mãos é realizado com apenas uma, ex: chuva) e inversão da mão dominante (HELOIR; GIBET, 2007).

Exemplos de variação estruturais são apresentados no quadro 9 abaixo.

Quadro 9 : variação distal proximal na representação de entonação por avatar.
 Fonte: Adaptado de Heloir; Gibet (2007, p.8)

Entonação	Imagem do avatar
Neutro	
Cansado	
Irritado	

Para a aplicação de estilo, modificadores de prosódicos podem ser adicionados em uma camada acima dos sinais que eles modificam (ADAMO-VILLANI et al., 2010).

3.5.3.8 Conclusões sobre síntese de sinais a partir de avatares

As iniciativas voltadas à síntese de sinais apresentam-se diretamente relacionadas à modelagem de um meio de representação das LS em meios digitais, às máquinas de tradução automática e à modelagem de avatares.

Cabe salientar que

Os vídeos em LS com intérpretes humanos aparentam ser a solução mais comum e aceita atualmente. As possibilidades de avatares de LS ainda são restritas e estão distantes de uma tradução automática (MÖBUS, 2010, p.574, tradução nossa).

Um dos aspectos críticos da síntese de sinais é o realismo da animação gerada. Esta qualidade em geral é prejudicada pela forma com que o avatar é modelado ou pelo modo como a animação é parametrizada. A aplicação de elementos auxiliares para definir aspectos prosódicos da mensagem pode aumentar a qualidade da sinalização.

A técnica de síntese de sinais também permite a redução da quantidade de dados para transmissão da mensagem em LS, contudo requer maior processamento de dados para sua realização. No caso da TV digital esta capacidade de processamento é limitado pelo hardware dos receptores.

A principal vantagem do uso de avatares é a possibilidade de troca de intérprete sem que o conteúdo tenha que ser modificado ou então a edição do conteúdo sem necessidade de regravação, como ocorreria caso o conteúdo em LS estivesse representado em vídeo com um intérprete real.

3.6 CARACTERIZAÇÃO DAS TECNOLOGIAS PARA PRODUÇÃO E APRESENTAÇÃO DE LÍNGUAS DE SINAIS

O quadro 10 apresentado a seguir descreve os principais métodos e técnicas identificados para a produção e apresentação de línguas de sinais segundo a taxonomia de Fajardo, Vigo e Salmeron (2009).

São caracterizados três grupos principais de tecnologias: as baseadas em escritas de sinais, o uso de vídeo com intérprete humano e o uso de avatares.

Quadro 10: Análise das tecnologias para produção de LS partir da taxonomia de Fajardo, Vigo e Salmeron (2009).

Fonte: os autores

Parâmetros	Gerais	Tecnologia	HamNoSys	Stokoe	SignWriting	Ahmed e Seong (2006)	AGA-Sign - Denardi (2006)	Debevc et al. (2009)	Segala (2010)	Phan, Nguyen e Bui (2009)	PoliLibras – Januário, Leite e Koga (2010)	VSigns Papadogiorgaki et al. (2004)	Visicast - Elliot et al (2007)
		Tarefa	Estudos linguísticos	Estudos linguísticos	Comunicação	Comunicação	Comunicação	Navegação	Aprendizagem	Comunicação	Comunicação	Comunicação	Comunicação
		Modalidade	Escrita	Escrita	Escrita	Escrita	Escrita	Gestual	Gestual	Gestual	Gestual	Gestual	Gestual
		Gravação	Orientada ao usuário	Pré-gravada	Pré-gravada	Orientada ao usuário	Orientada ao usuário	Orientada ao usuário	Orientada ao usuário				
		Aplicação em outros dispositivos	Sim	Sim	Sim	Sim	Sim	Não, exclusivo para web	Web e DVD	Não, exclusivo para PC	Não, exclusivo para PC	Não, web	Sim, web e TV
	Teste por usuários	Sim	Sim	Sim	Sim	Não	Sim	Sim	Não	Não	Sim	Sim	
	Comunicação Sinalizada	Espaço de Sinalização	-	-	-	-	Os definidos pelo SignWriting	Parte superior do corpo	Parte superior do corpo	Parte superior do corpo	Mãos	Parte superior do corpo	Parte superior do corpo
		Transmissão da LS	-	-	-	-	Arquivos de imagens animadas	Em separado	Em separado	-	-	Em separado	Em separado
		Outros	-	-	-	-	Animação de escrita de sinais	Aplicado na comunicação multimodal	Aplicado no curso Letras Libras	Uso de datilografia para sinais não encontrados no dicionário	Uso de notação proprietária para armazenar a animação	Uso de SignWriting para especificar a animação	Uso do HamNoSys para especificar a animação
	Avatares	Tipo de Avatar	-	-	-	-	Cartoon	-	-	Humano	Humano	Humano	Humano
		Dimensões	-	-	-	-	2D	-	-	3D	Semantróide 3D	3D	3D
		Localização da animação	-	-	-	-	Servidor ou cliente	-	-	Servidor	Servidor	Servidor	Servidor
		Localização da renderização	-	-	-	-	Servidor ou cliente	-	-	Cliente	Cliente	Cliente	Cliente
Tecnologia gráfica		-	-	-	-	GIF	-	-	Java3D, VRML e H-Anim	IK, X3D, Processing	VRML e H-Anim	VRML	

Continuação do quadro 10.

Parâmetros	Gerais	Tecnologia	WebSign – Jemni e Elghoul (2007)	SignSmith – Grieve e Smith (2001)	RybenáTV – Amarin et al.(2010)	Krapež e Solina, (1999)	Moemedi e Connan (2010)	Sport Sign – Othman, Ghoul e Jemni, (2010)	Tavares, Corradine e Breda (2005)	Xu et al. (2008)	Zwitsers et al. (2004)	San-Segundo et al. (2008)	
		Tarefa	Comunicação	Comunicação	Comunicação	Comunicação	Comunicação	Comunicação	Comunicação	Comunicação	Comunicação	Comunicação	Comunicação
		Modalidade	Gestual	Gestual	Gestual	Gestual	Gestual	Gestual	Gestual	Gestual	Gestual	Gestual	Gestual
		Gravação	Orientada ao usuário	Orientada ao usuário	Orientada ao usuário	Pré-gravado	Orientada ao usuário	Orientada ao usuário	Orientada ao usuário	Orientada ao usuário	Orientada ao usuário	Orientada ao usuário	Orientada ao usuário
		Aplicação em outros dispositivos	Não	Não	Web e TV	Computadores	Não	Sim, dispositivos fixos e móveis	Não	Sim	Não	Sim	Não
		Teste por usuários	Sim	Não	Sim	Não	Não	Sim	Sim	Sim	Não	Sim	Sim
	Comunicação Sinalizada	Espaço de Sinalização	Parte superior do corpo	Parte superior do corpo	Parte superior do corpo	Parte superior do corpo	Cabeça	Parte superior do corpo	Parte superior do corpo	Parte superior do corpo	Parte superior do corpo	Não incorpora aspectos não manuais	Parte superior do corpo
		Transmissão da LS	Web, em separado	Web, em separado	TV	CD	Vídeo	Vídeo	Em separado	Em separado	Em separado	Em separado	Em separado
		Outros	-	Uso de notação Stokoe para parametrizar a animação	Realiza a tradução da legenda oculta	Junção de vídeos pré-gravados	-	Uso de concatenação de vídeos gerados por avatar	Uso de concatenação de vídeos	Síntese de sinais a partir de base de dados de movimentos	Síntese de sinais a partir do reconhecimento de movimentos e notação Stokoe	Síntese de sinais a partir do reconhecimento de movimentos e notação Stokoe	Síntese de sinais a partir do reconhecimento de voz
	Avatares	Tipo de Avatar	Humano	Humano	Cartoon	-	Humano	Humano	-	Humano	Humano	Cartoon	
		Dimensões	3D	3D	2D	-	3D	3D	-	3D	3D	2D	
		Localização da animação	Servidor	Servidor	Servidor	-	Servidor	Servidor	-	Servidor	Servidor	Servidor	
		Localização da renderização	Cliente	Cliente	Servidor	-	Servidor	Servidor	-	Cliente	Servidor	Cliente	
		Tecnologia gráfica	H-Anim, X3D	VRML, Web3D	GIF	-	-	FLV, XML	-	H-Anim, VRML	-	-	

3.7 CONCLUSÕES SOBRE ARTEFATOS DE APOIO À ACESSIBILIDADE E AMBIENTES DE ACESSO

Diante a necessidade em se produzir recursos que permitam o acesso ao conteúdo audiovisual por públicos surdos, a literatura demonstrou que diferentes abordagens podem ser utilizadas. Pode-se afirmar que as legendas textuais e a legenda com vídeo em língua de sinais são os principais artefatos identificados, cada uma com diretrizes específicas para sua produção.

O processo de tradução está intimamente associado à produção destes artefatos, sendo a tradução intersemiótica a modalidade presente tanto na produção de legendas em língua oral quanto de legendas em línguas de sinais.

Para a produção de legendas textuais na língua oral, o procedimento de legendagem sugere a adoção de legendas com múltiplos níveis de condensação do texto, de modo que cada espectador possa escolher o nível ao que se sente mais confortável para a leitura. A taxonomia de Teixidor (2008) permite caracterizar as legendas textuais, definindo, entre outras características, a intenção de seu uso, o modo de sua produção e sua forma de entrega.

Para a legendagem com línguas de sinais é possível o uso de intérpretes, humanos ou avatares, ou então o uso de um sistema de escrita de sinais. As tecnologias que podem ser utilizadas na produção de legendas em LS são classificadas e caracterizadas diante a taxonomia de Fajardo, Vigo e Salmeron (2009), que apresentam a notação escrita como uma forma de representação das LS, assim como o vídeo tradicional e o uso de avatares 3D ou 2D.

O vídeo é o meio mais utilizado atualmente para a comunicação em LS, porém possui aspectos que dificultam sua produção em escala. A codificação do vídeo em LS, que visa reduzir a quantidade de dados necessária, deve respeitar os requisitos decorrentes da comunicação por sinais e privilegiando as áreas da face durante a codificação.

A produção de legendas com vídeo de LS por meio de avatares demonstra-se promissora em relação à otimização de recursos utilizados na produção, tais como a possibilidade de edição de conteúdo e a flexibilização nas modalidades de entrega da legenda.

Do mesmo modo, o uso de sistemas de escrita de sinais permite sistemas automatizados ou mais eficientes que a gravação de vídeo em termos de demanda de dados e possibilidade de edição.

A possibilidade de legendas em escrita de sinais pode ser identificada diante o uso desta modalidade de escrita em aplicativos e

conteúdos na web, entretanto os processos associados a este tipo de artefatos não estão claramente definidos e seu uso integrado aos outros tipos de legendas é investigado no próximo capítulo.

A tradução automática apresentada diante as diferentes modalidades de tradução pode utilizar desde técnicas de reconhecimento de voz na produção de legendas textuais, o reconhecimento de vídeo em LS para geração de conteúdo escrito em LS, a junção de vídeos pré-gravados, e o uso de diferentes tipos de avatares para apresentação da LS.

Os dicionários e glossários bilíngues apoiam os processos de produção e permitem o acesso dos surdos às mensagens e termos desconhecidos, desenvolvendo sua própria cultura quando associado às LS e interagindo com a cultura ouvinte.

O próximo capítulo descreve a integração dos métodos para produção de legendas em um conjunto de processos relacionados, os quais oferecem uma arquitetura reutilizável para a concepção de sistemas voltados à produção de artefatos de apoio a acessibilidade dos surdos ao audiovisual.

4. PROPOSTA DE MODELO DE REFERÊNCIA

O modelo de referência é elaborado a partir dos enunciados teóricos e experiências práticas descritas no capítulo anterior, ele faz o mapeamento dos possíveis processos para a produção de artefatos de acesso dos surdos e destina-se à integração de iniciativas relacionadas ao desenvolvimento de conteúdo acessível aos surdos.

Os processos e artefatos que constituem o modelo de referência possibilitam o projeto de sistemas para desenvolvimento de recursos de apoio à acessibilidade de surdos ao audiovisual por meio da escolha de alternativas identificadas, sendo que o caminho entre os possíveis processos é decorrente do artefato que se deseja ofertar e dos recursos técnicos e humanos disponíveis.

O modelo informa também os requisitos que devem ser atendidos e funcionalidades que devem estar disponíveis nos sistemas de apoio à produção de artefatos para acessibilidade.

Como requisitos para potencializar o acesso dos surdos ao audiovisual por meio do modelo proposto estão a adequação do tamanho das legendas a diferentes públicos surdos, o uso de texto escrito e sua adequação ao tempo de leitura de diferentes usuários, e o uso de vídeo com LS.

A possibilidade de uso em cenários com conteúdo on-line ou pré-gravado depende dos processos que serão utilizados, porém o foco está em ambientes com conteúdo pré-gravado, pois não são consideradas as restrições em termos de atraso na produção de legendas.

Como benefício em relação aos trabalhos correlatos está a integração entre os desenvolvimentos existentes e a identificação de lacunas a serem preenchidas por pesquisas futuras.

Os modelos são descritos a partir da *Unified Modelling Language* (UML) por meio de diagramas de casos de uso, classes do domínio e sequências de interação relacionadas à produção do audiovisual acessível.

Em complemento é apresentado um diagrama que mapeia e inter-relaciona as iniciativas voltadas à acessibilidade do surdo e um modelo de processos e artefatos que visa integrar os processos identificados.

A validação do modelo é realizada por meio do método Delphi com consulta aos especialistas a respeito das diretrizes e recomendações derivadas do modelo. Os questionários utilizados e respectivos resultados são apresentados e analisados no capítulo cinco.

4.1 MODELOS DE CASOS DE USO PARA PRODUÇÃO E ACESSO AO AUDIOVISUAL

Um caso de uso descreve as principais interações que usuários realizam com o sistema para atingir determinado objetivo. Os diagramas de casos de uso relacionam os possíveis casos de uso diante o modelo para produção de artefatos de acessibilidade dos surdos.

A intenção dos diagramas é apresentar os possíveis atores e as tarefas que estes desempenham em cenários com produção dos artefatos identificados no modelo.

Os casos de uso do modelo de referência proposto estão divididos em dois grupos principais: os relacionados à produção dos artefatos para acessibilidade, na figura 79; e os relacionados à fruição do audiovisual acessível, apresentados na figura 80.

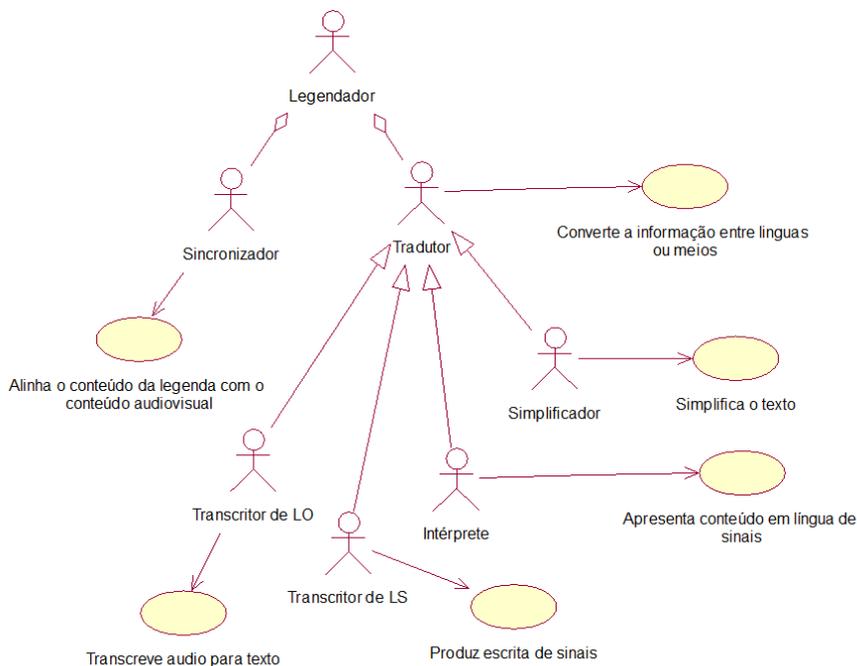


Figura 79: Modelo de casos de uso para a produção de artefatos de apoio.

Durante a produção de artefatos de apoio, são identificados 7 tipos de atores: o legendador, o tradutor, o transcritor de língua oral (LO), o transcritor de LS, o simplificador, o intérprete e o sincronizador.

O tradutor é um tipo de ator que pode possuir diferentes especializações, atuando tanto na tradução interlinguística, intersemiótica ou intralinguística.

O papel do simplificador é gerar legendas com múltiplos níveis de redução do texto com o objetivo de oferecer opções para usuários com diferentes níveis de leitura, realizando a tradução intralinguística.

Por sua vez o intérprete de LS apresenta o conteúdo traduzido em língua de sinais.

O sincronizador faz o alinhamento temporal entre a legenda e o conteúdo do audiovisual.

O transcritor de LO é responsável pela transcrição literal do áudio para o texto, como no caso do *closed-captions*. Já o transcritor de LS codifica os sinais apresentados por um intérprete de LS.

O legendador agrega os papéis desempenhados pelo sincronizador e pelo tradutor.

Todas as atividades relacionadas realizam determinados tipos de tradução e os papéis dos atores podem ser desempenhados tanto por agentes humanos ou então de modo automático.

O diagrama da figura 80 apresenta os casos de uso que o surdo pode desempenhar para acessar o conteúdo apresentado pelo audiovisual.

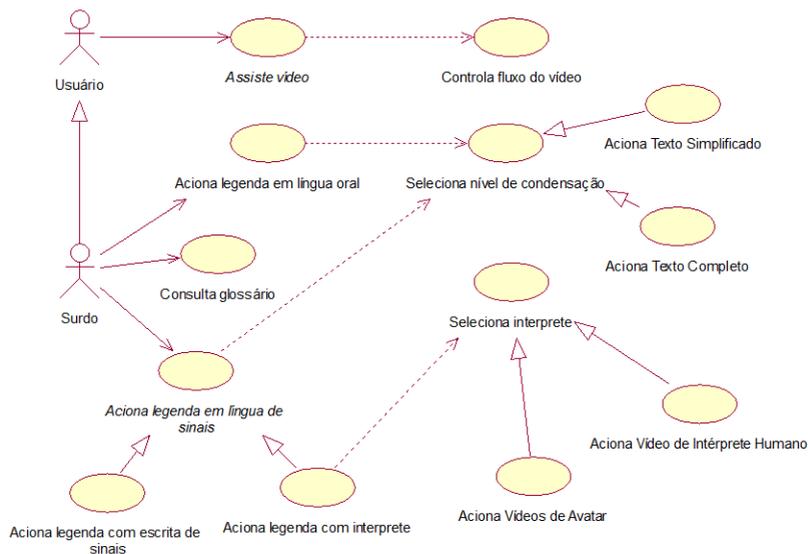


Figura 80: Modelo de casos de uso para o acesso do surdo ao audiovisual.

O surdo é um usuário do audiovisual que pode ativar a exibição de legenda em LS ou em língua oral. A legenda em língua oral é sempre apresentada na modalidade escrita, já a legenda em LS pode ser apresentada por meio do vídeo de um intérprete, real ou virtual, ou por meio de um sistema de escrita.

A consulta a um glossário pode ocorrer eventualmente caso o usuário entre em contato com um sinal inédito ou desconhecido.

Ao ativar a legenda escrita é possível selecionar o nível de condensação do texto, além disso, podem ser definidos outros parâmetros para sua apresentação, como cor e tamanho da fonte.

Por meio dos diagramas apresentados é possível relacionar os papéis e funcionalidades de sistemas voltados à acessibilidade de surdos ao audiovisual.

O próximo tópico descreve os modelos de classes de entidades que estão relacionados à realização dos casos de uso.

4.2 MODELOS DE CLASSES

Os diagramas de classe representam modelos sobre as entidades existentes e seus relacionamentos no contexto do desenvolvimento dos artefatos para acessibilidade. São apresentados modelos voltados às classes de artefatos, às técnicas de produção de conteúdo em LS e aos glossários.

4.2.1 Modelo de classes de artefatos para acessibilidade

A figura 81 apresenta as classes de artefatos que podem ser produzidos para promover a acessibilidade do audiovisual para surdos.

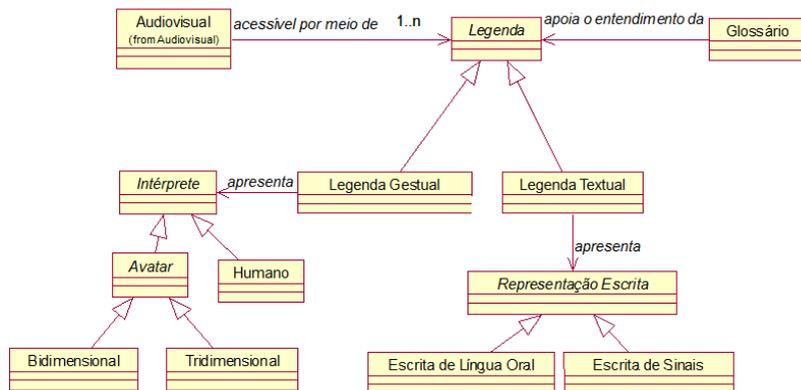


Figura 81: Hierarquia de classes de artefatos para acessibilidade ao audiovisual.

A legenda é o elemento que possibilita o acesso do surdo ao audiovisual, ela pode se dividir nos tipos gestual ou textual.

No caso de legendas gestuais, o agente que apresenta a mensagem é um intérprete humano ou avatar. No caso de avatares, pode ser utilizado um avatar tridimensional e realista ou um avatar bidimensional e com maior abstração para interpretação do conteúdo.

Já no caso de legendas textuais, um sistema de escrita é utilizado, o qual pode se basear na língua oral, como a do português ou inglês, ou então em um sistema de escrita voltado às línguas de sinais, como o EliS ou o SignWriting.

Todos os tipos de legendas são acessíveis por meio de um glossário bilíngue, que descreve os termos utilizados na tradução do conteúdo.

O próximo tópico apresenta com mais detalhes a hierarquia de classes associadas às técnicas para produção de legendas gestuais.

4.2.2 Modelo de Técnicas de Produção de Legendas Gestuais

As legendas gestuais em LS podem ser desenvolvidas a partir de técnicas manuais e automáticas, conforme figura 82.

As técnicas manuais utilizam os métodos *Slideshow*, *Serial Relay* ou *Parallel Relay* (FELS et al., 2009).

As técnicas automáticas de concatenação de vídeos geram legendas em LS a partir da junção de filmagens de um intérprete real, previamente realizadas por meio de técnicas manuais.

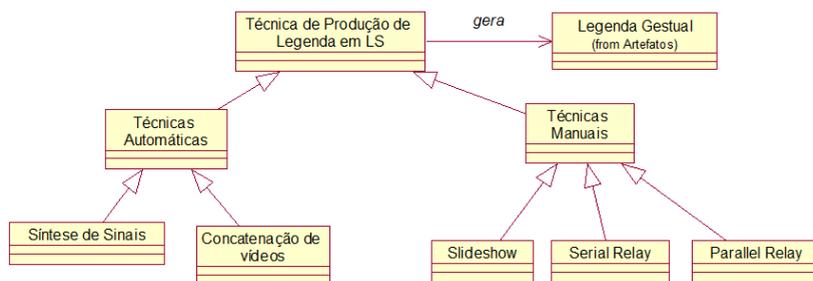


Figura 82: Tipos de técnicas para produção de legendas com LS.

As técnicas automáticas de síntese de sinais geram conteúdo em LS por meio de avatares. O modelo de classes para síntese de sinais é descrito na figura 83, que contém as entidades relacionadas à síntese de sinais por meio da animação de avatares.

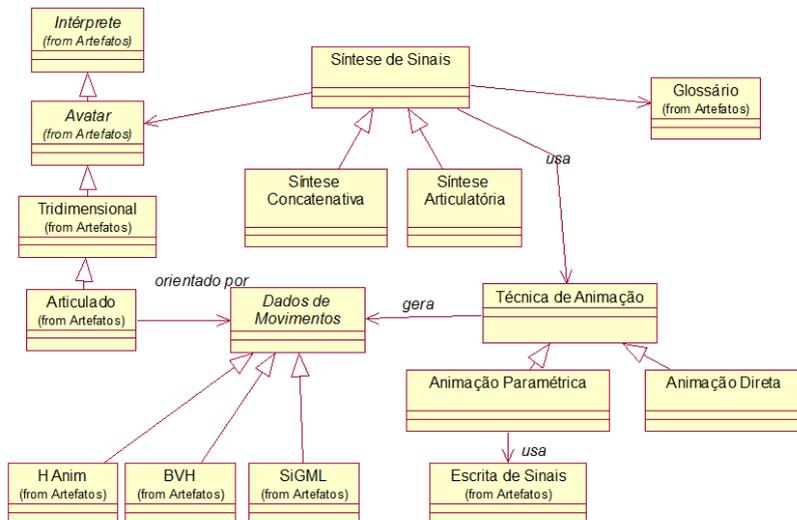


Figura 83: Hierarquia de classes para animação de avatares.

A técnica para síntese de sinais pode ser concatenativa ou articulatória e atua gerando dados de movimentos para um avatar articulado.

A síntese concatenativa faz a junção de dados de movimentos pré-gravados. A definição dos dados de movimentos referentes a um sinal pode ser feito com a animação direta do avatar ou por captura de movimentos.

A síntese articulatória usa a animação paramétrica baseada em elementos fonológicos das LS representados por meio da escrita de sinais.

Os dados de movimentos gerados pelas técnicas de animação podem ser armazenados no glossário, em diferentes padrões de dados, tais como o BVH, H-Anim e SIGML.

O glossário pode armazenar este tipo de dados e tem sua hierarquia descrita no próximo tópico.

4.2.3 Modelo de classes para o glossário

A figura 84 apresenta as classes que associam as legendas aos glossários. Esta associação é feita por meio de sentenças compostas por palavras representadas em um ou mais sistemas de escrita. Além disso cada termo armazenado no glossário pode ser representado por uma descrição fonética, um vídeo ou então por dados de movimentos que serão realizados por um avatar.

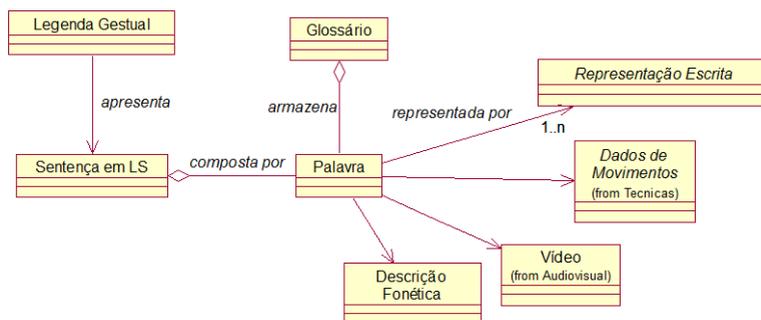


Figura 84: Diagrama de classes do glossário.

O glossário contém as descrições das palavras utilizadas nas legendas, podendo também armazenar dados para sua animação por um avatar. O glossário pode também armazenar metadados que indiquem a regionalidade dos termos usados da legenda.

A próxima sessão descreve o modelo de atividades associadas aos processo de produção de artefatos para acessibilidade de surdos ao audiovisual.

4.3 MODELOS DE ATIVIDADES DOS PROCESSOS PARA PRODUÇÃO DE ARTEFATOS

Esta sessão descreve a modelagem dos processos que podem ser realizadas para produzir os artefatos de acessibilidade. Os modelos de atividades estão inicialmente divididos em grupos de processos relacionados as:

- legendas de texto oral;
- legendas de língua de sinais em vídeo com intérprete humano;
- legendas de línguas de sinais com intérpretes virtuais, e;
- legendas em escrita de sinais.

Posteriormente estes modelos de atividades são apresentados de forma integrada em um modelo de processos e artefatos.

4.3.1 Modelo de atividades para produção de legendas em texto oral

O modelo apresentado na figura 85 descreve os passos para a produção de legendas textuais na língua oral.

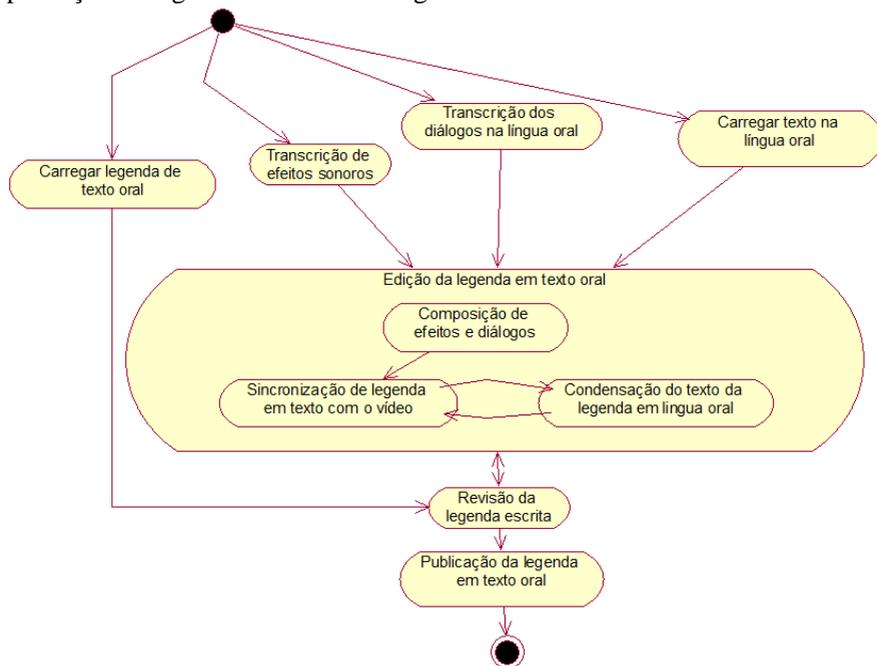


Figura 85: Sequência de atividades para produção de legendas em texto.

O passo inicial para a produção de legendas em texto é a transcrição dos diálogos e dos efeitos sonoros. Para este trabalho, efeitos sonoros englobam a música, ruídos e outros sons que não as falas.

Alternativamente é possível o carregamento de um arquivo de legendas pré-existente.

A edição da legenda em texto oral permite que elementos sejam corrigidos ou então frases sejam unidas ou separadas, de modo a evitar sobrecarga de texto da tela.

Durante a edição, a legenda é sincronizada com o conteúdo do vídeo principal do audiovisual.

Em complemento é possível sua condensação, respeitando-se o limite de caracteres adequado para a velocidade de leitura e gerando versões diferentes de legendas para o mesmo audiovisual.

A revisão verifica se o sincronismo e a quantidade de caracteres estão adequados.

O passo seguinte é a publicação da legenda, seja ela para o ambiente *web* ou de TVD, em um modelo de empacotamento compatível com as plataformas de entrega.

4.3.2 Modelo de Atividades para Produção de Legendas de Sinais com Intérprete Real

A interpretação das línguas de sinais para geração de legendas em LS pode ser feita diretamente a partir do áudio do audiovisual ou então a partir da sua transcrição.

Para a tradução para Libras é necessária a análise prévia do conteúdo a ser traduzido, e neste momento podem ser definidos os termos utilizados na tradução.

A filmagem do intérprete utiliza uma das técnicas manuais de produção de vídeo em LS e gera um vídeo que pode ser editado para sincronização ou adição de efeitos.

Após sua edição o vídeo é codificado e revisado. Na revisão, os termos utilizados nas legendas devem ser validados pela sua presença no glossário e então o vídeo gravado é publicado.

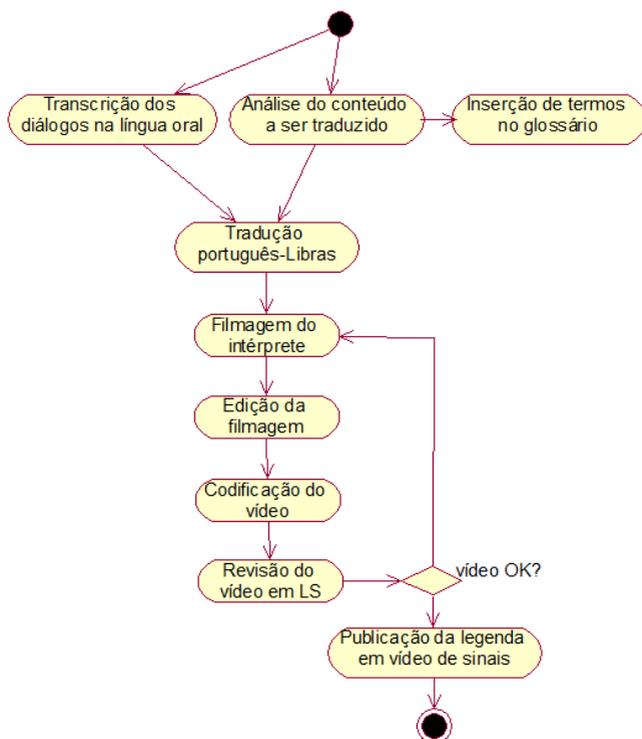


Figura 86: Sequência de atividades para produção de legenda com vídeo.

4.3.3 Modelo de Atividades para Produção de Legendas em Escrita de Sinais

A oferta de legendas em escrita de sinais depende da realização das etapas descritas no diagrama da figura 87. O desenvolvimento deste tipo de legenda inicia pela tradução do conteúdo audiovisual para a Libras, que pode gerar tanto um texto em escrita das LS, ou então um vídeo com um intérprete em LS, que passa pela transcrição e gera a legenda em escrita de sinais. Como terceira alternativa é possível o carregamento de um texto em escrita de sinais pré-elaborado.

Após sua elaboração, o texto é editado, sendo sincronizado com a linha temporal das mensagens do conteúdo principal, podendo eventualmente ser condensado ou simplificado. A condensação ocorre de modo a permitir a exibição simultânea dos caracteres na tela em tempo hábil para a leitura do usuário. Após sua edição, a legenda é revisada e publicada.

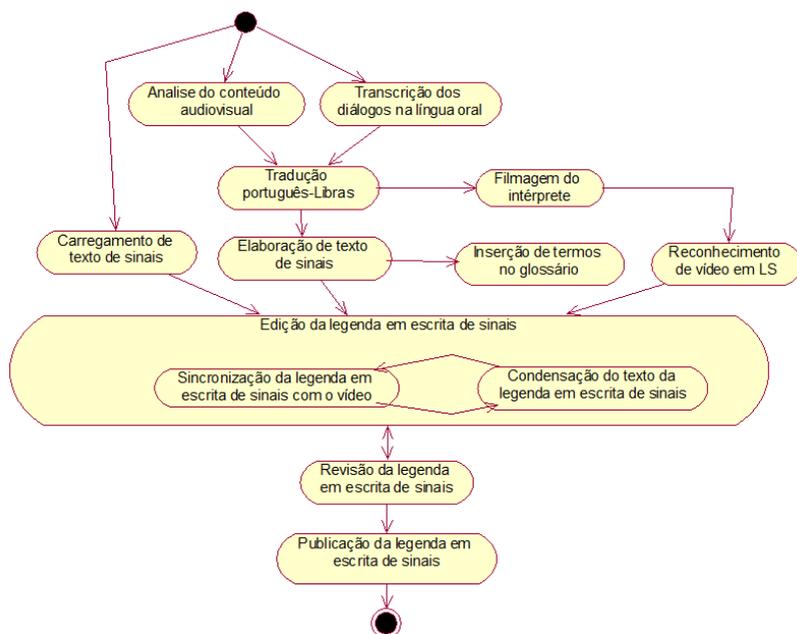


Figura 87: Atividades para produção de legendas em escrita de sinais.

4.3.4 Modelo de atividades para produção de Legendas em Língua de Sinais com Intérprete Virtual

O uso de intérpretes virtuais, ou avatares, na apresentação das línguas de sinais pode oferecer ganhos de escala na produção de conteúdo acessível, visto que estão associados aos sistemas de tradução automática, apesar de poderem ser utilizados de forma independente de sistema de tradução.

Considerando-se aspectos referentes à qualidade da tradução automática e da animação apresentada, a aplicação de síntese de sinais por meio de avatares pode ser aplicada de modo a promover a acessibilidade em conteúdos audiovisuais em plataformas digitais.

A figura 88 apresenta as possíveis atividades realizadas no processo de produção de legendas em LS com intérprete virtual.

O processo inicia pela elaboração de um texto de entrada para um sistema de tradução automática. Este texto pode ser em língua oral ou em LS, e deve passar por um processo de sincronização com o conteúdo do vídeo principal.

O passo seguinte é gerar os parâmetros de animação a partir do texto de entrada, estes parâmetros são fornecidos a um sistema de animação que gera os quadros-chaves referentes aos movimentos e aplica as poses necessárias à sinalização do conteúdo. É na definição dos parâmetros de animação que podem ser aplicados os elementos prosódicos, os quais modificam o tempo de cada fase dos sinais realizados, de modo a tornar a animação menos robótica.

Após a geração dos parâmetros de animação, estes são aplicados em um avatar compatível com os parâmetros gerados. O sistema de animação realiza os sinais e renderiza um vídeo com o conteúdo em LS que deve ser codificado priorizando as áreas de interesse para comunicação por sinais.

Na revisão são analisadas a qualidade e sincronia da animação gerada, caso estes aspectos não permitirem a compreensão do conteúdo traduzido os parâmetros de animação precisam ser redefinidos.

Com a aprovação do vídeo, a legenda em LS pode ser publicada.

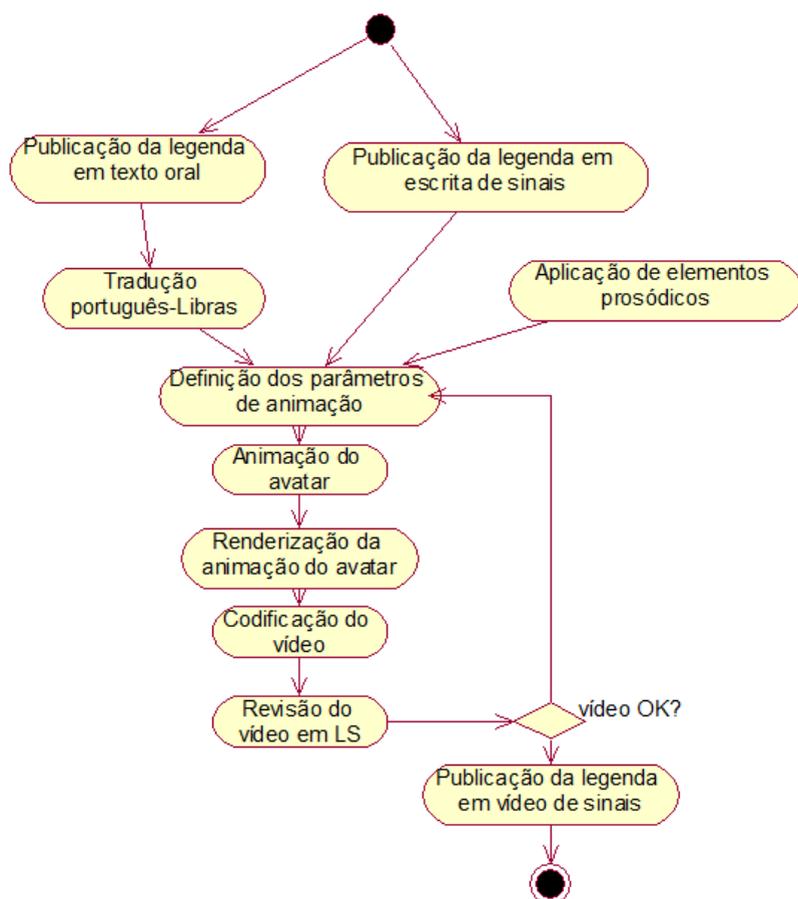


Figura 88: Atividades para produção de legenda de LS com avatares.

4.3.5 Modelo Integrado das Atividades para Produção de Artefatos de Apoio ao Acesso dos Surdos ao Audiovisual

As alternativas de processos de produção de artefatos de acessibilidade são apresentados no diagrama da figura 90.

O elemento inicial na produção de legendas é a audiovisual, principalmente a sua informação sonora e que gera um texto na língua oral ou orienta a tradução do conteúdo para LS.

O glossário está associado à transcrição e tradução dos conteúdos, pois registra e fornece informações sobre os termos utilizados nestes processos.

As legendas textuais podem ser geradas a partir da transcrição textual da informação sonora do audiovisual e sua sincronização, podendo haver sua simplificação para geração de legendas de múltiplos níveis. O áudio pode ser extraído tanto pelo seu reconhecimento por máquina quanto por sua transcrição manual. O uso exclusivo da informação sonora na tradução pode levar a resultados de menor qualidade em relação à processos que considerem também a informação visual das cenas traduzidas.

A partir das legendas textuais pode-se partir para a gravação manual de vídeo em LS, ou então utilizar processos de tradução automática, gerando legendas em LS.

A tradução automática do texto em português para a Libras gera uma sequência de sinais representado por meio da escrita de sinais ou então gera os parâmetros para a animação de um avatar. A tradução automática também pode gerar sequências de pequenos vídeos que podem ser concatenadas em um vídeo contínuo.

A legenda em LS com intérprete real pode ser produzida tanto a partir da tradução e interpretação direta do audiovisual, quanto pela tradução e interpretação das legendas. Além disso pode ser feita a tradução da legenda oral escrita para uma descrição escrita dos sinais e sua interpretação posterior.

A transcrição do conteúdo do vídeo da legenda em LS para um texto em escrita de sinais pode ocorrer tanto por processo manual ou pelo reconhecimento automático. O texto em escrita de sinais também pode ser gerado a partir da tradução manual a partir da informação do audiovisual. Este texto escrito é fragmentado e sincronizado, eventualmente simplificado, gerando a legenda em escrita de sinais.

A legenda em escrita de sinais, associada à uma descrição fonética dos sinais, pode orientar a geração de dados para a gesticulação do avatar. Aspectos prosódicos como estilo de apresentação, variações

estruturais e entonação dos sinais podem ser definidos durante a edição da escrita de sinais e servem como parâmetros para animação, que produz os dados de movimentos dos sinais.

Uma vez gerados os dados para a animação do avatar, este passa por um processo de renderização, de modo a transformar os modelos tridimensionais em uma sequência de imagens que pode ser visto pelo usuário. É na renderização que o vídeo é codificado e os parâmetros de qualidade para visualização das LS devem ser respeitados.

Como itens entregáveis ao usuário de modo a gerar e possibilitar o acesso ao conteúdo audiovisual traduzido estão a própria legenda em texto oral, podendo esta ter múltiplos níveis de condensação de conteúdo, e a legenda em escrita de sinais, que são acessíveis por meio de um glossário bilíngue.

Com relação aos artefatos de acessibilidade com intérprete de LS, podem ser entregues o vídeo com intérprete real ou com intérprete virtual, codificados de acordo com os critérios de legibilidade de conteúdo em LS. Outro elemento entregável são os dados de movimentos, que podem ser transmitidos ao receptor do usuário e serem renderizados após a recepção. Além disso os glossários podem ser considerados artefatos entregáveis, pois estes auxiliam no entendimento das legendas.

A legenda utilizada para apresentação do modelo é apresentada na figura 89. Processos são representado por retângulos, e artefatos são representados por losangos. Artefatos efetivamente enviados ao usuário estão preenchidos em cinza.

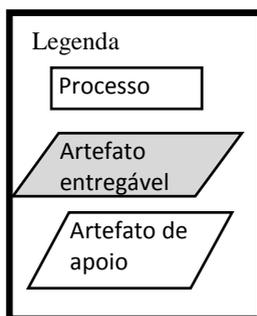


Figura 89: Legenda para modelo de alternativas de processos para produção de legendas.

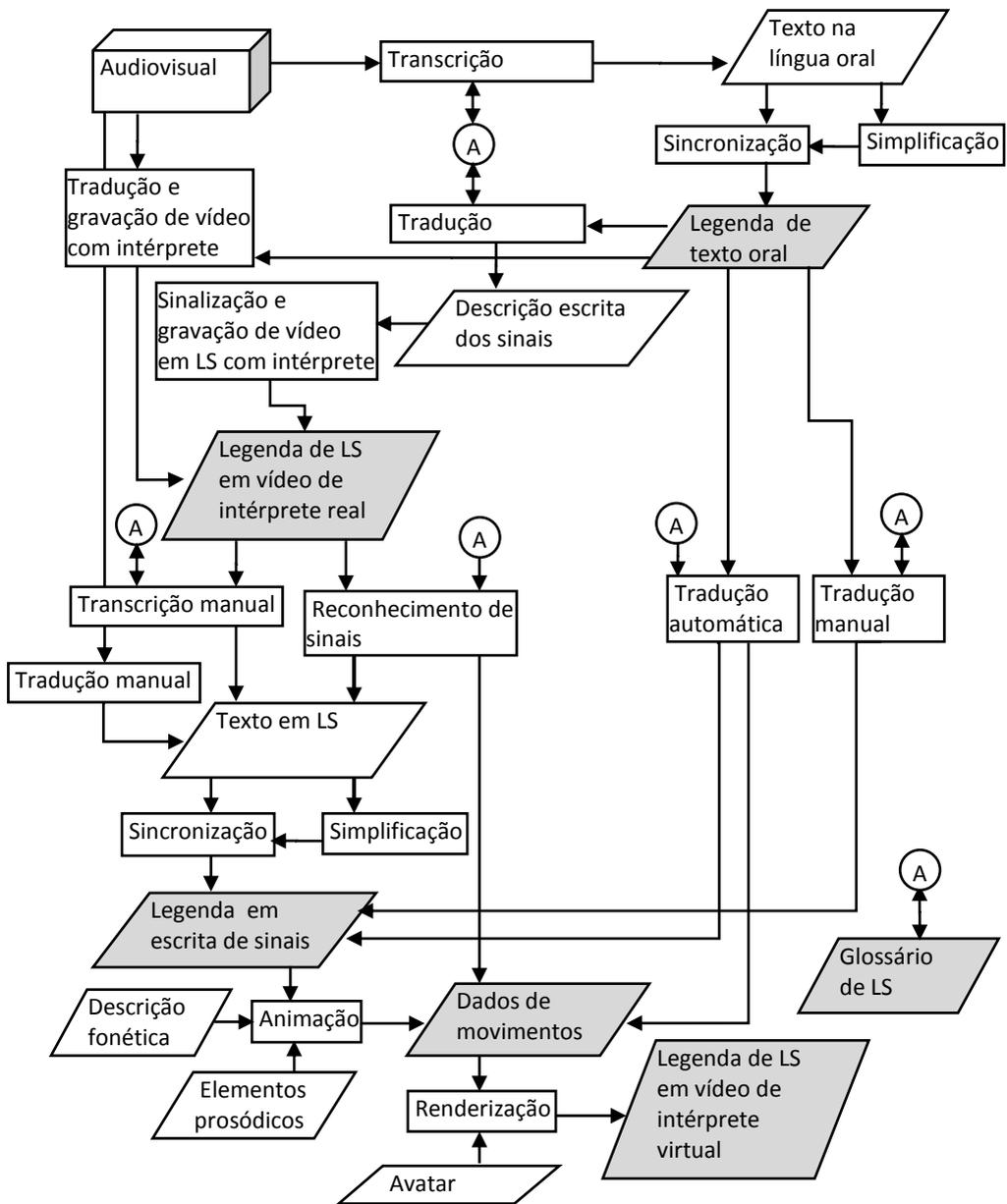


Figura 90: Modelo de alternativas de processos e artefatos para produção de legendas

As iniciativas identificadas para a produção de artefatos podem ser associadas à diferentes modalidades de comunicação nos cenários com audiovisual acessível aos surdos.

O diagrama da figura 91 apresenta estas associações. As modalidades de comunicação estão representadas nos nós da rede apresentada, cada ligação é identificada por uma letra e representa um grupo de procedimentos e abordagens para tradução.

A produção de artefatos pode ser realizada, por exemplo, a partir da língua oral falada do audiovisual e sua transcrição em legendas na língua oral, as quais podem ser traduzidas com o uso de um intérprete virtual para a LS (A-G). Ou os artefatos podem ser produzidos a partir do audiovisual, gerando-se uma legenda em língua de sinais com um interprete real, que é convertida em um texto em escrita de sinais (C-E).

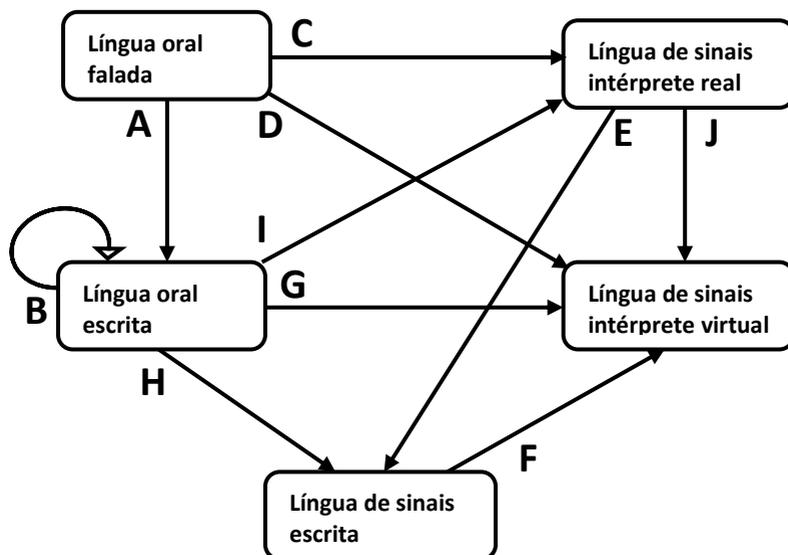


Figura 91: Modalidades de comunicação e possíveis traduções para desenvolvimento de audiovisual acessível aos surdos.

O quadro 11 enumera as iniciativas que apoiam os processos identificados no modelo.

Quadro 11: Mapeamento de iniciativas relacionadas ao modelo

A	Geração de legendas em	Lambourne et al., 2004; Bain; Bason;
---	------------------------	--------------------------------------

	texto a partir do reconhecimento de voz e da tradução manual.	Wald, 2002; Lu, 2010; Wald, 2006; Rander; Looms, 2010; Tveit, 2004; Burnham et al., 2008; Neves, 2005; 2007.
B	Simplificador de textos.	Caseli et al., 2009; Candido et al., 2009.
C	Tradução e interpretação de LS a partir das informações do audiovisual.	Nascimento, 2010; 2011; JISC, 2010; Oi; Tang; Wai, 2009.
D	Reconhecimento de áudio do audiovisual e tradução para LS com avatar.	San-Segundo et al., 2008.
E	Reconhecimento de LS, por meio de visão computacional ou por luvas de dados.	Efthimiou et al., 2009; Bedregal; Dimuro; Costa, 2006; Efthimiou et al., 2009; Ghoul; Jemmi, 2010; Dreuw et al., 2010.
F	Conversão de escrita de sinais para animação de intérprete virtual de LS	Efthimiou et al., 2009; Denardi, 2006; Papadogiorgaki; Grammalidis; Makris, 2005; Kennaway; Glauert e Zwitterlood, 2007; Zwiterslood et al., 2004; Verlinden; Zwitterlood; Frowein 2005.
G	Animação de avatar em LS a partir de texto em língua oral	Karpouzis et al., 2007; Xu et al., 2008; Phan; Nguyen; Bui, 2009; Kaneko et al., 2010; Amaral; De Martino; Angare, 2011.
H	Tradução com dicionário de escrita de sinais	Sign Puddle, 2012; Souza, 2005b.
I	Tradução e interpretação da LS a partir de texto	Santos; Oliveira; Oliveira, 2010; Segala, 2010; FELS et al., 2009; JISC, 2010; Tavares; Corradine; Breda, 2005.
J	Síntese de intérprete virtual a partir da gravação de intérprete real ou o uso de intérprete real para modelagem de animações	Maldonado, 2007; Adamo-Villani; Benil, 2005.

4.4 DESCRIÇÃO DO MODELO DE ACORDO COM A VISÃO DE MUNDO

A visão de mundo a partir da teoria da complexidade permite descrever o modelo de referência proposto como sendo um conjunto de diferentes artefatos e estratégias que permitem a produção de artefatos de apoio ao acesso de agentes surdo ao espaço conceitual delimitado pelo audiovisual produzido por agentes ouvintes.

Este acesso é promovido quando oferecidos artefatos que permitam o desenrolar de estratégias dos surdos de forma a superarem as barreiras estabelecidas pela língua utilizada para a apresentação do conteúdo do audiovisual.

Os tradutores constituem agentes que pertencem tanto ao espaço conceitual dos surdos quanto dos ouvintes, e que realizam a integração entre estes dois espaços.

4.5 CONSIDERAÇÕES SOBRE O MODELO

O modelo apresentado por meio de diferentes diagramas descreve os tipos de artefatos de acessibilidade e as atividades associadas aos processos para sua produção.

Os casos de uso descrevem os papéis desempenhados pelos atores envolvidos na produção das legendas e as atividades que estes desempenham.

Os diagramas de classe apresentam a hierarquia existente entre os artefatos os relacionam aos formatos de mídia e técnicas utilizadas.

Os diagramas de atividades descrevem as etapas para os processos de produção de legendas escritas e em vídeo.

As abordagens identificadas no terceiro capítulo são inter-relacionadas, permitindo mapear as possibilidades de uso integrado, que são apresentadas na figura 90.

Os próximos capítulos descrevem a validação do modelo diante um grupo de especialistas, uma implementação de referência e sua aplicação em cenários de uso.

5. QUESTIONÁRIO DELPHI

Este capítulo descreve a validação do modelo diante a consulta à comunidade de especialistas por meio de um questionário aplicado segundo o método Delphi.

O questionário busca levantar recomendações sobre:

- Aspectos sobre o público surdo e a acessibilidade ao audiovisual;
- Os artefatos que promovem a acessibilidade de surdos;
- Papéis e responsabilidades nos processos de produção no contexto da produção destes artefatos.

Foram enumeradas assertivas que fundamentam um modelo para produção de legendas e estas foram divididas em grupos de perguntas que englobam aspectos sobre surdos e audiovisual, os papéis desempenhados na produção de legendas, os artefatos de apoio, e o processos de produção de legendas em português, em LS em vídeo, com avatares e em escrita de sinais.

A busca aos especialistas aconteceu por meio da plataforma Lattes. Foi utilizada a chave de busca “surdos + televisão”, “surdos + legendas” e “surdos+ audiovisual”, e onde foram enviados convites para responder ao questionário aos 30 pesquisadores mais relevantes em cada uma das chaves inseridas. A caracterização dos especialistas convidados a responder o questionário é apresentada no anexo D, onde nota-se predominância de formação nas áreas de letras, comunicação e educação.

Houve no total cerca de 90 especialistas consultados, dos quais 17 responderam na primeira rodada e 8 na segunda. Esta quantidade de respostas foi suficiente para validar as recomendações com um nível de significância de 5%.

A cada proposição o especialista consultado pode concordar ou não concordar, atribuindo um grau de importância para a recomendação. Sendo que todas as respostas eram optativas, podendo o especialista não enviar seu parecer caso não se sentisse familiarizado com a questão apresentada. Ao final de cada grupo de proposições foi disponibilizado um campo para comentários.

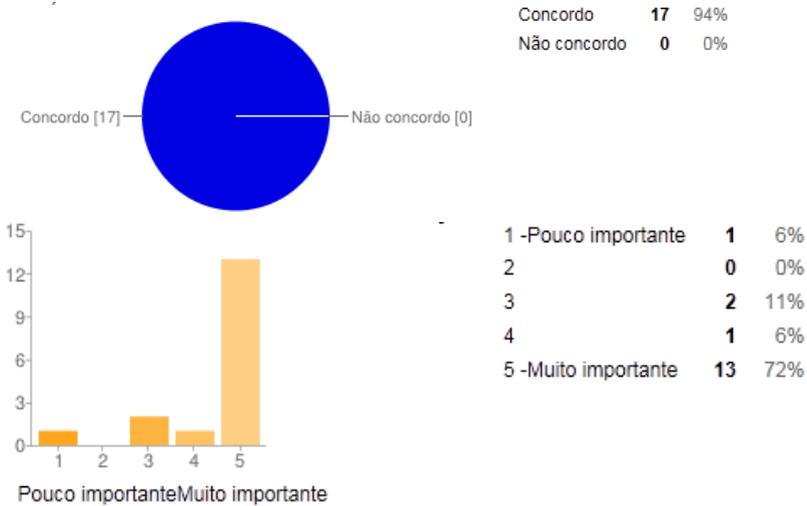
As assertivas estão agrupadas em cinco grupos e os resultados e opiniões dos especialistas a seu respeito são apresentados a seguir.

Os itens com menos de 75% de aprovação foram revisados a partir dos comentários da primeira rodada e foram revalidados em uma nova rodada de consulta.

5.1 SURDOS E AUDIOVISUAL

Estas recomendações visam caracterizar as possibilidades de recursos utilizados pelos surdos no acesso ao audiovisual:

1. O acesso dos surdos ao audiovisual pode ser ampliado pela oferta de legendas escritas;



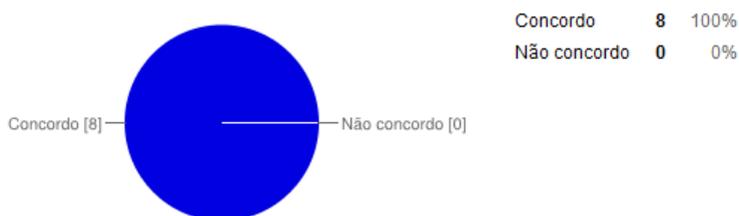
2. O acesso dos surdos ao audiovisual pode ser ampliado pela oferta de legendas em línguas de sinais;



3. A simplificação do texto da legenda amplia o acesso do surdo às mensagens do audiovisual;



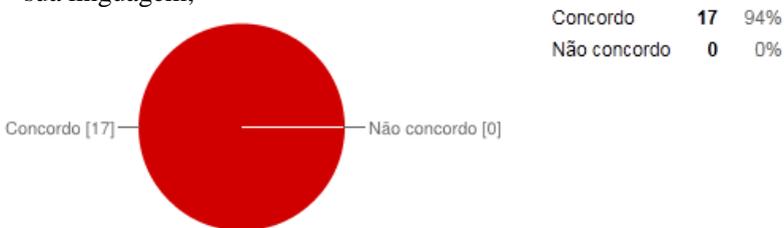
(2ª rodada) A simplificação do texto da legenda amplia o acesso do surdo às mensagens do audiovisual, porém também deve se oferecer acesso à legendas mais completas de modo a evitar restrições de vocabulário.

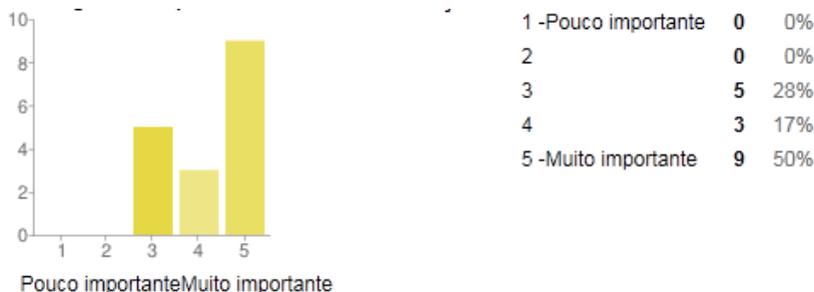


4. Deve ser permitida ao surdo a escolher o tipo de legenda que deseja utilizar, sendo que esta pode conter uma transcrição literal da fala, uma versão simplificada do texto, um intérprete de língua de sinais ou então legendas em escrita de sinais com os sinais apresentados pelo intérprete.



5. Um glossário ou dicionário bilíngue, que pode ser acessado para esclarecer termos usados nas legendas, potencializa o acesso do surdo ao audiovisual ao mesmo tempo em que permite desenvolver sua linguagem;





Comentários sobre artefatos

- "A questão 3 seria concordo parcialmente, dependo do que se refere a texto simplificado. Se for o uso de palavras mais conhecidas, de sentenças mais simples e diretas pode ser eficaz. No entanto, fazer em forma de glosas, como se fosse a escrita/nomeação dos sinais em glosas em português considero muito arriscado. Quanto a questão 2, a legenda é algo importante, pois há muitas pessoas ensurdecidas que se beneficiam da legenda e os surdos usuários da libras que estão em processo de aprendizado do português a podem utilizar como acesso a contextos diversificados da escrita do português, mas no momento histórico o uso exclusivo de legenda escrita priva os surdos do acesso à informação."
- "A apropriação da modalidade de linguagem sinalizada é imprescindível para o surdo (estou considerando a condição de surdez neurossensorial profunda bilateral e/ou surdez neurossensorial severa p/ profunda bilateral) se apropriar dos conhecimentos culturalmente veiculados, mas a apropriação do português escrito também é importante para diferentes possibilidades de inclusão em diferentes âmbitos (social, educacional, profissional...), de forma que a simplificação de conteúdos da modalidade escrita pode induzir à compreensão superficial de tais conteúdos, reduzindo as possibilidade de significação tb nessa modalidade de linguagem, frequentemente e infelizmente, reduzida a expressões de cumprimentos, classificações (numérica, categorial)."
- "Fiquei com duvida nessa questão: O acesso dos surdos ao audiovisual pode ser ampliado pela oferta de legendas em línguas de sinais; Legendas em língua de sinais? no formato escrito? No Brasil

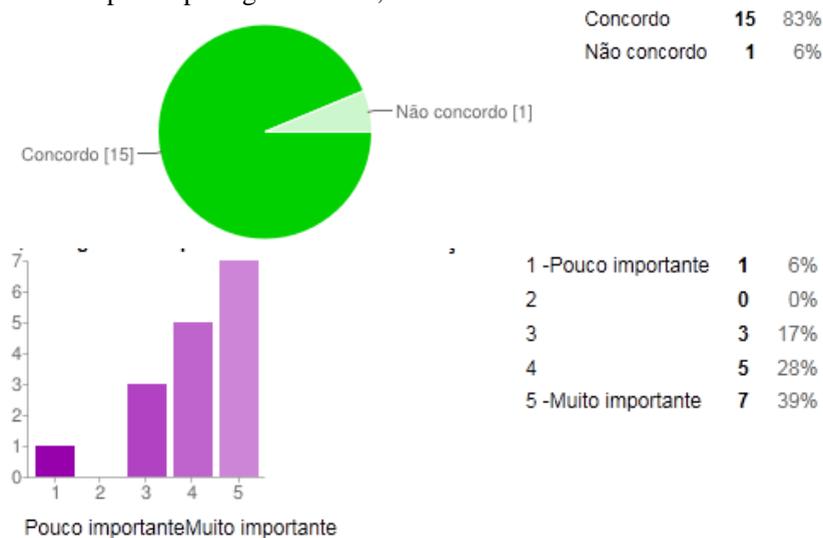
há duas escritas de sinais que estão em amplo debate, Sign Writing - mais difundida entre a comunidade surda e ouvinte, e Elis - menos difundida..como seria esse processo de implementação destas legendas escritas?"

- “A simplificação do texto como única opção tende a criar um espectador mais desinformado do que outros.”
- “Não concordo com a legenda em língua de sinais SE o programa for para adultos, porque já devem ter necessária proficiência em língua portuguesa, como cidadãos brasileiros.Para crianças surdas sim, deveria ser em, língua de sinais sinalizada. Nunca em escrita de sinais, pois considero uma maneira a mais de diferenciar a comunicação, dificultando-a, portanto.”
- “Parabéns pelo tema de pesquisa! Fiquei muito contente em ver que tem pesquisadores interessados em investigar as condições de acessibilidade comunicacional de pessoas surdas frente as normativas brasileira.”
- "Eu não concordo com a simplificação das legendas, pois o surdo fica com um vocabulário restrito se não tiver acesso a uma linguagem real. É importante o uso de legendas curtas e lentas, pois são mais facilmente lidas do que as longas e rápidas. também não concordo que a legenda possa ser escrita de qualquer forma, pois se o surdo aprender a língua escrita do país onde vive, ele poderá ter mais acesso ao mercado de trabalho e ao mundo acadêmico."

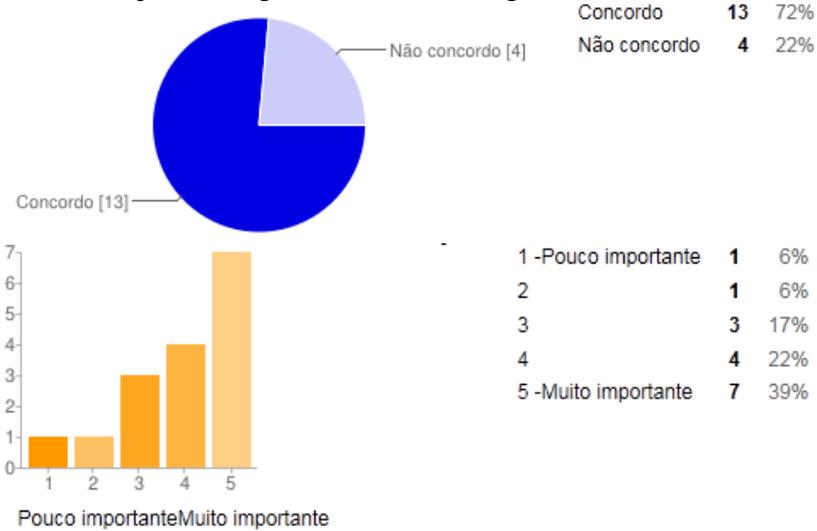
5.2 PAPÉIS

O objetivo destas questões é validar os nomes dos papéis desempenhados e suas responsabilidades nos processos de tradução e legendagem.

6. O transcritor de língua oral é responsável pela transcrição dos diálogos e efeitos sonoros das cenas do audiovisual, do português oral para o português escrito;



7. O transcritor de língua de sinais gera uma legenda em escrita de sinais a partir da legenda de vídeo em língua de sinais;



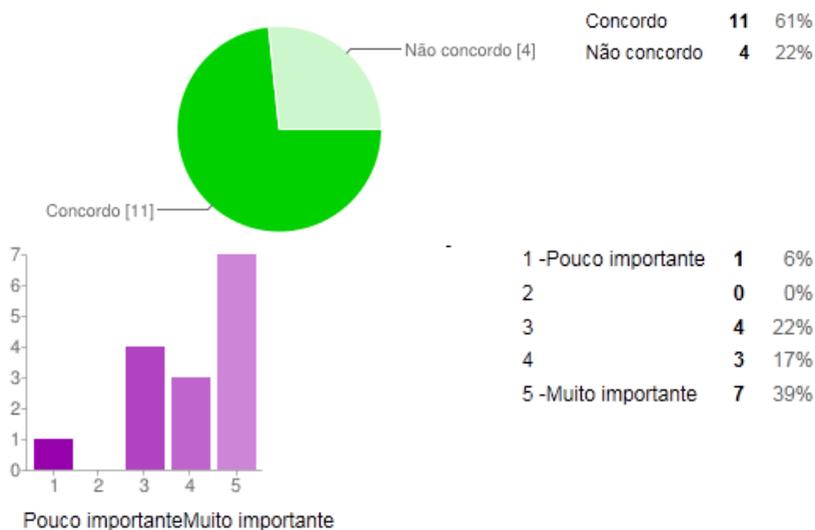
(2ª rodada) O transcritor de língua de sinais é um tradutor que gera um texto em escrita de sinais a partir de um vídeo em língua de sinais;



8. O intérprete é responsável pela apresentação do conteúdo em língua de sinais;



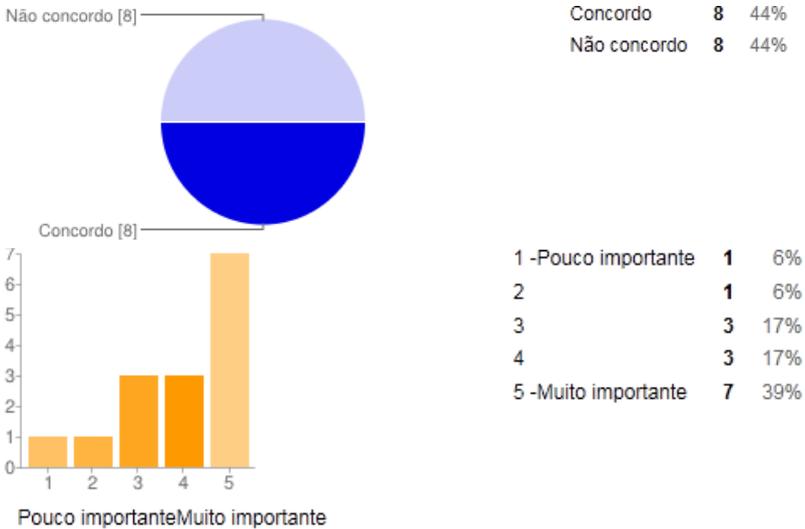
9. A redução da quantidade de texto exibida nas legendas, de modo a facilitar o entendimento é realizada pelo simplificador;



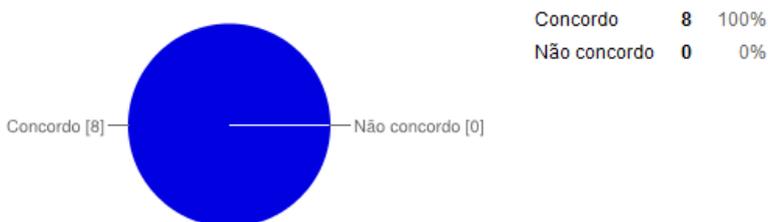
(2ª rodada) A redução da quantidade de texto exibida nas legendas, de modo a facilitar o entendimento é realizada pelo tradutor ao utilizar estratégias de condensação do texto;



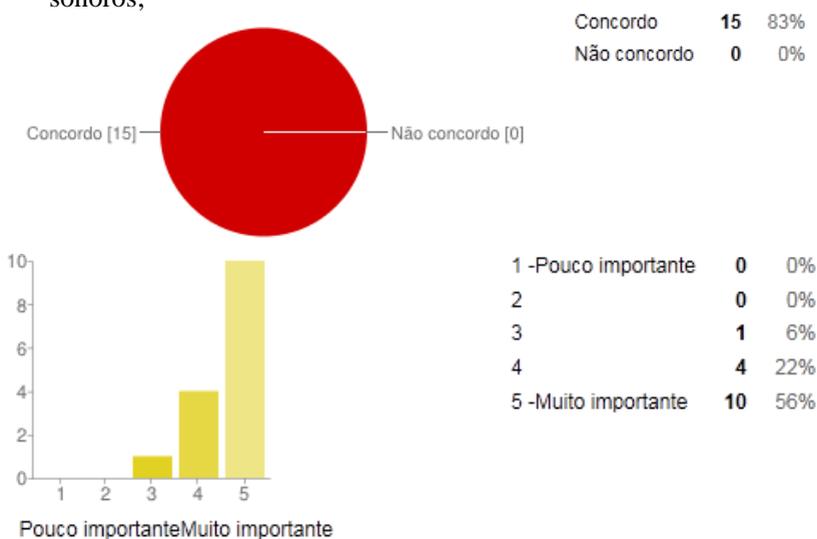
10. O tradutor converte a informação de uma língua para outra, podendo atuar também como transcritor, simplificador ou intérprete;



(2ª rodada) O tradutor converte a informação de uma língua para outra, sendo aconselhável que tenha conhecimentos sobre o tema abordado no conteúdo traduzido, atuando na transcrição, condensação ou interpretação;



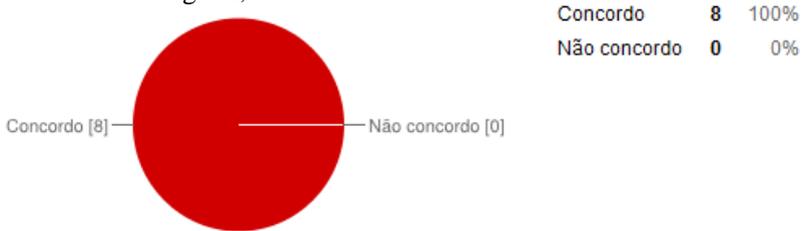
11. O sincronizador é responsável por alinhar o tempo de exibição da legenda com o tempo das falas das personagens e dos efeitos sonoros;



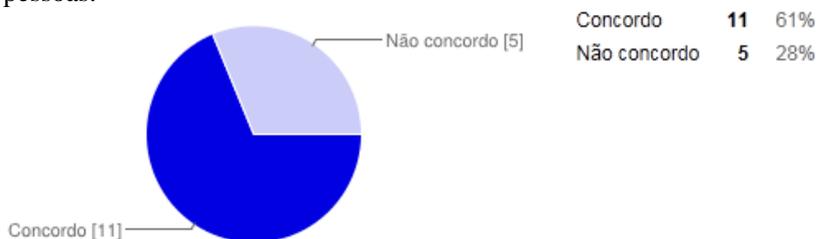
12. O legendador agrega os papéis do tradutor e do sincronizador;



(2ª rodada) O legendador atua em conjunto com o tradutor e é responsável pela sincronização do conteúdo traduzido com o conteúdo original;



13. Todos os papéis podem ser desempenhados por uma ou mais pessoas.



(2ª rodada) Todos os papéis podem ser desempenhados por uma ou mais pessoas, em geral uma atuando na tradução e outra na legendagem. O resultado da atividade destes papéis é validado por um revisor de conteúdo.



Comentários sobre papéis

- Para atuar como transcritor, intérprete e principalmente como simplificador é preciso ter uma especialização para se adequar às diferentes estratégias e objetivos de comunicação propostas. O tradutor pode atuar em parceria com o legendador. Fazendo o

trabalho em dois profissionais, pois usualmente o tradutor (pelo menos de Libras) desconhece os softwares e sistemas de legendagem.

- Ronnie, estás perguntando, afirmando, é esse o conceito? queres saber se o nome desses profissionais leva ao entendimento depois explicado? Fora o intérprete, os outros desconheço.
- Na transcrição muito se perde dos aspectos de paralinguagem. A expressão orofacial, por exemplo, muitas vezes é crucial p/ a interpretação adequada do sinal. Ainda, existe legislação específica p/ a formação e atuação do intérprete. Não é simplesmente alguém fluente em Libras, não basta só isso. Numa sala de aula, por exemplo, precisa de um domínio mínimo dos conteúdos trabalhados p/ não interferir na compreensão acerca do conteúdo interpretado. Uma mesma pessoa executar todas as tarefas listadas no questionário simplifica a relevância do papel do intérprete que, infelizmente e principalmente no contexto educacional, não tem formação adequada...
- Em minha opinião a pergunta 6 invalida a 7 e a 8. Mas como fiz a ressalva que se o programa for infantil é importante que haja a língua de sinais, nesse caso a resposta à pergunta 8 passa a ser válida. Mas para adultos, não. Defendo a resposta dada na 6. Você está propondo cinco profissões para o exercício desse trabalho (intérprete, transcritor, simplificador, legendador e sincronizador)? É demais. Bastaria o intérprete e o sincronizador, já que esta é uma esfera técnica diferente.
- "No meu entendimento o trabalho do tradutor e do intérprete se diferenciam, portanto, em função da falta de familiaridade com esta área de formação deste profissional - vinculado ao curso de letras, creio que todas as funções aqui explicitadas devem ter características peculiares ao exercício de suas atividades, possibilitando aos usuários da LIBRAS uma melhor qualificação dos serviços mencionados. Parabéns, mais uma vez!!!"
- Considero muito importante que mais de uma pessoa trabalhe na legendagem, pois o papel de uma máquina pode gerar tradução muito errônea. Sinto que uma equipe deve cuidar de todos os passos, que podem ser realizados com a ajuda de maquinário específico, mas deve ter supervisão de profissionais entendidos na área de surdez e de comunicação/ recursos técnicos/ eletrônicos.

- (2ª rodada) na tradução para a Libras a partir do português há muitas informações visuais dependentes de conhecimento prévio de mundo para a correta seleção dos sinais.

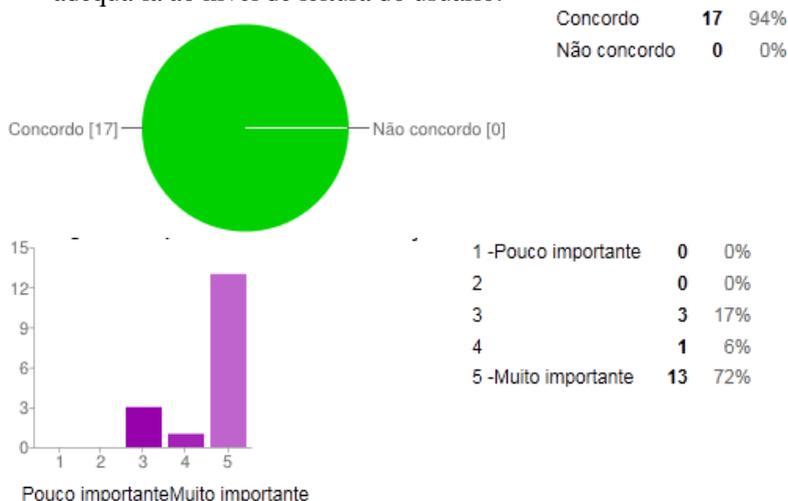
5.3 ARTEFATOS

As seguintes recomendações caracterizam os artefatos de apoio ao acesso dos surdos ao audiovisual:

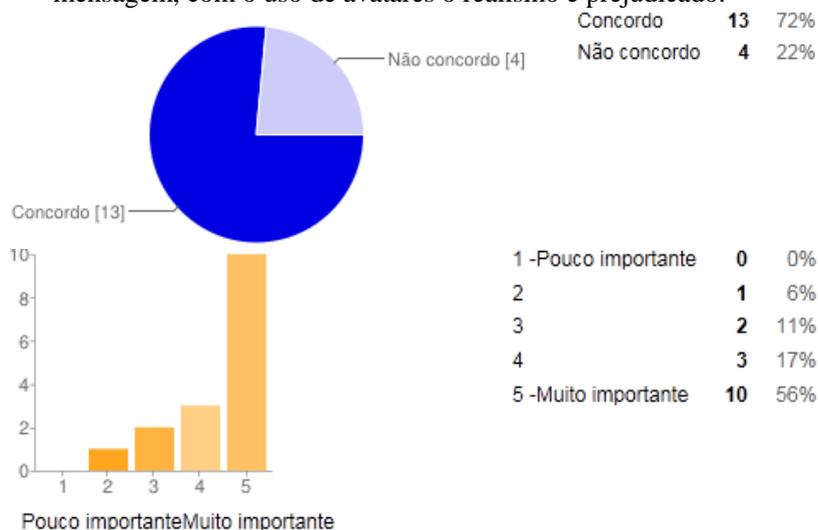
14. A legenda pode ser entendida como o elemento que apoia o entendimento da mensagem do vídeo, podendo se considerar exemplos de legendas: a legenda em texto; em vídeo com língua de sinais; e em escrita de sinais.



15. A legenda em português escrito pode ter diferentes níveis de condensação, desde uma transcrição literal da mensagem até um resumo adequado ao tempo de leitura. Deve ser possível a escolha do nível de condensação do conteúdo da legenda escrita, de modo a adequá-la ao nível de leitura do usuário.



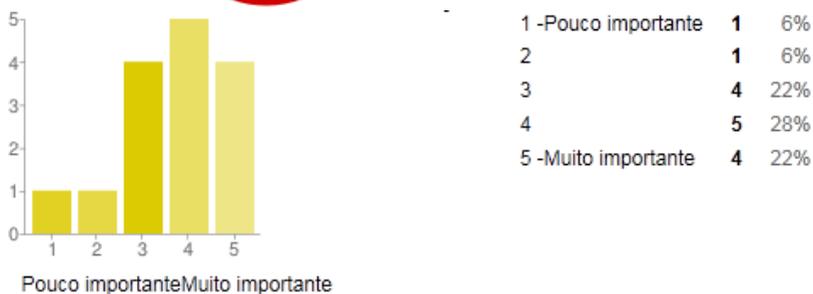
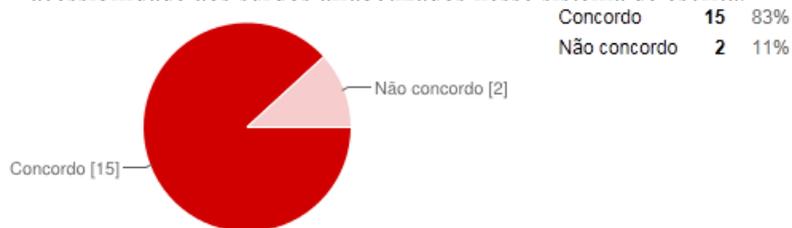
16. A legenda em língua de sinais pode utilizar intérpretes humanos ou avatares. Intérpretes humanos conferem maior realismo à mensagem, com o uso de avatares o realismo é prejudicado.



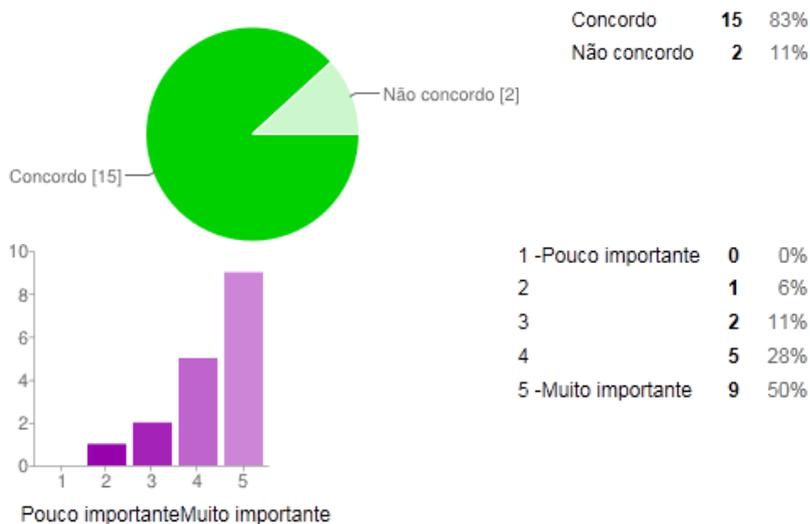
(2ª rodada) A legenda em língua de sinais pode utilizar intérpretes humanos ou avatares. Intérpretes humanos conferem maior realismo à mensagem, e o uso de avatares está associado às técnicas automáticas.



17. A oferta de legendas em escrita de sinais pode promover a acessibilidade aos surdos alfabetizados nesse sistema de escrita.



18. É vantajosa a oferta simultânea de legendas em português quanto em línguas de sinais;



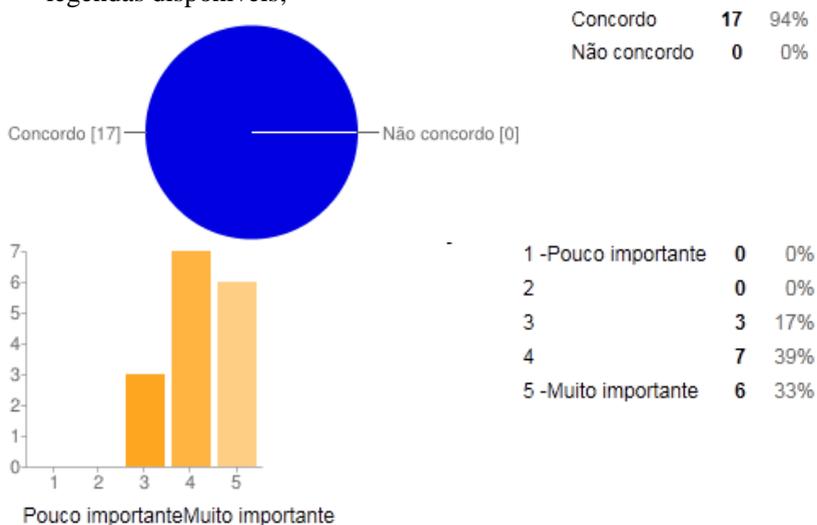
Comentários sobre Artefatos de Apoio

- Acredito que o avanço da tecnologia permite que os avatares atendam satisfatoriamente aos objetivos de comunicação com legenda em Libras.
- “Não acredito em análise homogeneizada da escrita. Cd um tem singularidades próprias no enfrentamento do processo de apropriação dessa modalidade de linguagem. De forma que não acredito em níveis, etapas, fases universalizadas, mas penso ser importante um conhecimento mínimo acerca do perfil do público-alvo p/ quem será destinada a legenda, a fim de adequá-la ao repertório, vocabulário, conhecimentos semânticos...
- É importante que o surdo aprenda português. Ter diferentes níveis de complexidade da legenda pode encarecer bastante o produto. Mas também pode ser um recurso bom para se aprender português, principalmente se o programa puder ser revisto.
- Na questão 18, penso que a possibilidade de o surdo efetuar a escolha por uma ou outra tipo de legenda seja mais adequada, para não atrapalhar o acesso a informação, visto que os sistemas gesto-visual e escrito são diferentes, por esse motivo discordo de utilizá-los simultaneamente. Garantir a escolha é um respeito as condições linguísticas dos surdos que poderão optar, conforme seu perfil linguístico e necessidades às modalidades disponíveis.
- Acho importante a opção de legenda em português escrito, com níveis diferenciados de acordo com a escolaridade, também acho importante vídeos em Língua de Sinais ou escrita de sinais, mas não acho bom que aconteçam simultaneamente. Dessa forma, pode-se apresentar legendas em momentos diferentes para os surdos, de acordo com o seu desempenho em leitura e em língua de sinais.

5.4 PRODUÇÃO DE LEGENDA EM PORTUGUÊS ESCRITO

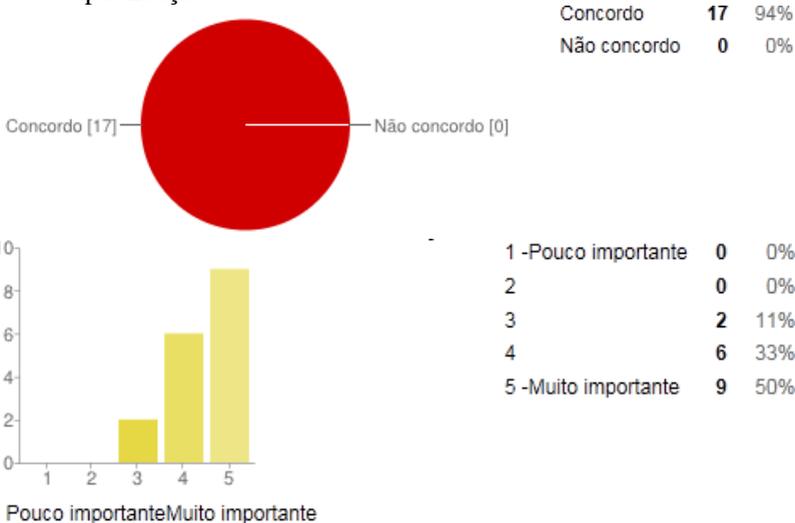
Estas questões tratam dos elementos e etapas envolvidas na produção de legendas em português escrito

19. As legendas em português escrito podem ser geradas a partir da informação sonora do audiovisual, ou a partir de outras formas de legendas disponíveis;

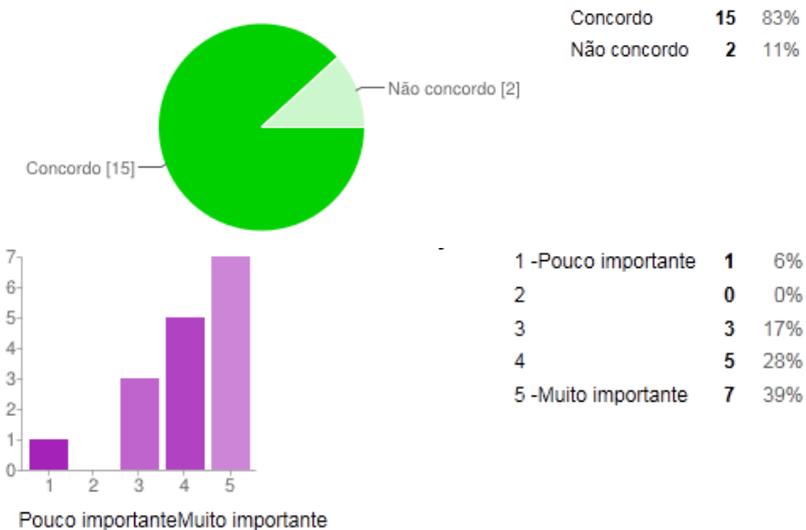


20. Os passos para a produção de legendas escritas em português são:

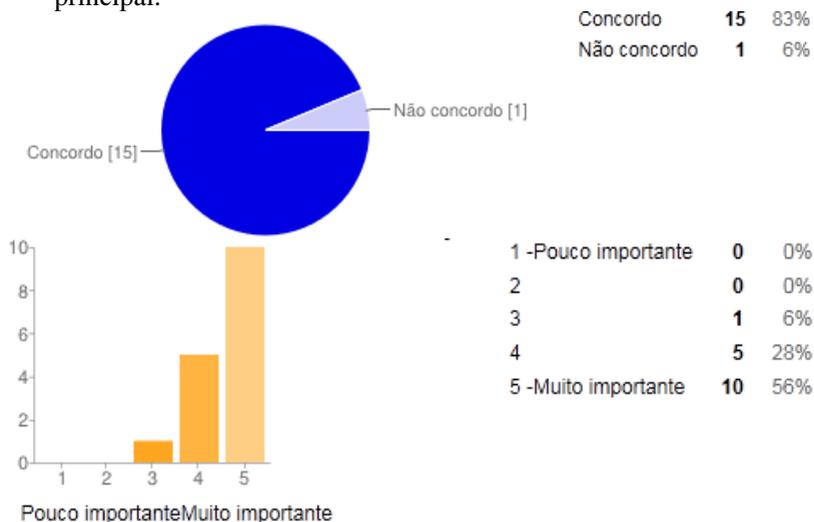
- transcrição dos diálogos e efeitos sonoros;
- edição, que inclui a composição, sincronização e condensação;
- revisão;
- publicação.



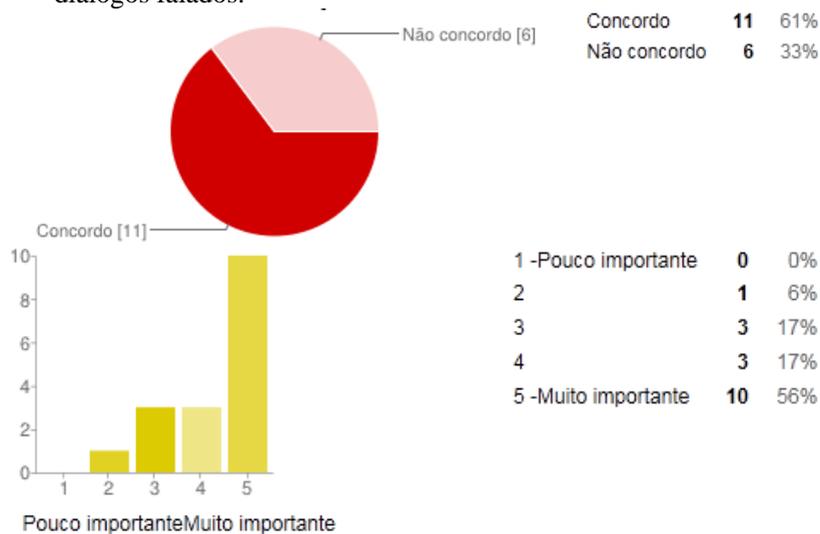
21. Quando houver uma legenda em língua oral disponível, os passos são revisão (caso houver necessidade ocorre a edição) e publicação.



22. A simplificação/condensação do texto das legendas escritas deve ocorrer de modo que seja mantido o sincronismo com o vídeo principal.



23. As legendas em português escrito podem ser produzidas pelo reconhecimento automático de áudio, ou pela digitação manual dos diálogos falados.



(2ª rodada) As legendas em português escrito podem ser produzidas pela digitação manual dos diálogos falados, ou pelo reconhecimento automático de áudio, que neste caso deve ser seguido de um processo de revisão para evitar distorções linguísticas.



Comentários sobre Produção de Legendas em Português Escrito

- Penso ser importante disponibilizar todas as formas de legenda, mas não ao mesmo tempo. Tal vez fosse mais apropriado o surdo ter a opção de escolha, no menu, do tipo de legenda que deseja acessar.
- Sou pela digitação manual, pois o reconhecimento automático de áudio pode gerar distorções linguísticas, se é que estou entendendo corretamente.
- "questão 21 - desconheço o processo de tradução da língua oral por isso fiquei na dúvida se seria apenas os dois processos mencionados. Questão 23 - creio que a legenda digitada automaticamente não garantirá a qualidade dos conteúdos das interlocuções, uma vez que dependerá da interpretação do mediador/tradutor e não do domínio de uma simples técnica de transposição de uma língua para outra. Dito de outro modo, nas relações dialógicas o tradutor levará em conta diferentes aspectos para efetuar o trabalho, tais como: contexto cultural dos falantes; aspectos linguísticos - relacionados ao estudo as áreas da fonologia, morfologia, semântica, sintaxe e pragmática tanto da fala como da escrita; aspectos extralinguísticos da mensagem como por exemplo: levar em conta as emoções, sentimentos dos interlocutores, que na maioria das vezes estão relacionadas com entoação, pausa e/ou silêncio dos falantes na língua oral - devem ser considerados nesse trabalho. Pelos motivos explicitados, acredito que a garantia desses aspectos na elaboração das legendas de uma modalidade para outra sempre dependerá do profissional/tradutor."

muitas vezes encontra-se erros grosseiros quando a tradução é feita diretamente do áudio, sem nenhuma revisão.

5.5 PRODUÇÃO DE LEGENDA DE VÍDEO EM LÍNGUA DE SINAIS

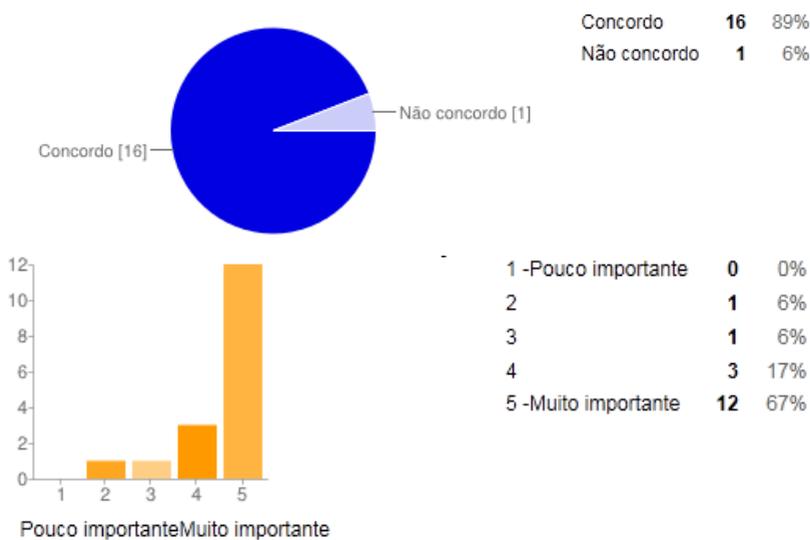
Estas questões tratam dos elementos e etapas envolvidas na produção de legendas em vídeo de LS.

24. As legendas de vídeo em Libras podem ser geradas a partir da informação sonora do audiovisual, ou a partir de outras formas de legendas disponíveis;

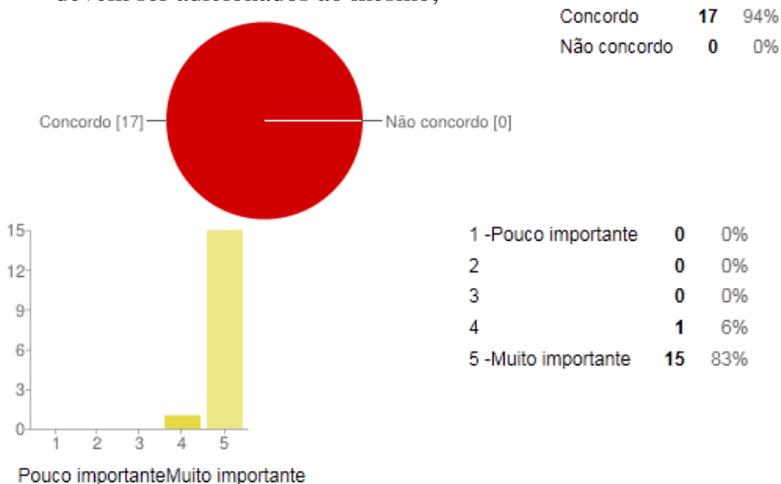


25. Os passos para produção de legenda de vídeo em Libras com intérprete real são:

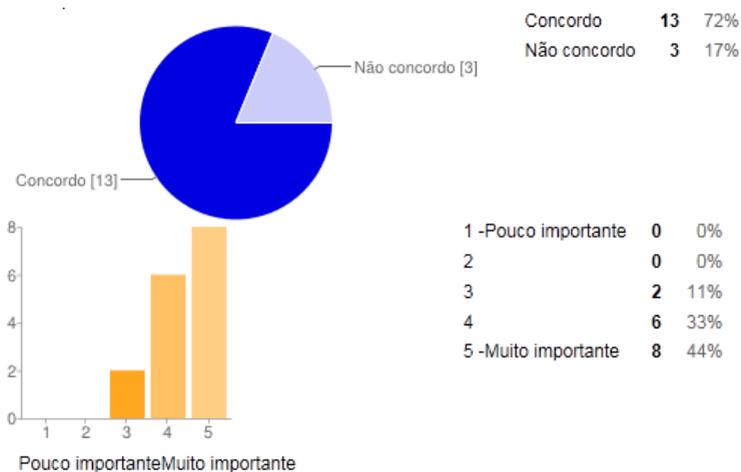
- Análise do conteúdo a ser traduzido
- Tradução português para Libras
- Filmagem do Intérprete;
- Edição da filmagem;
- Codificação do vídeo;
- Revisão;
- Publicação.



26. O tradutor e intérprete de LS deve ter contato prévio com o conteúdo audiovisual a ser traduzido, de modo a estabelecer um corpo de linguagem e definir os termos que serão utilizados nas legendas em LS. Termos inéditos, ou que não existam no glossário, devem ser adicionados ao mesmo;



27. As técnicas manuais para produção de conteúdo em Libras baseiam-se na gravação de vídeos;



(2ª rodada) A produção de conteúdo em Libras baseia-se principalmente na gravação de vídeos.



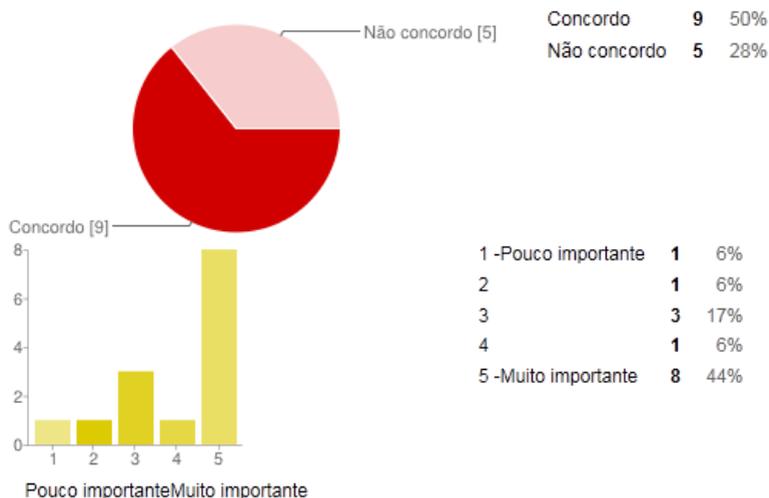
Comentários sobre produção de legenda em LS

- A inserção de sinais e expressões inéditas deve ocorrer mediante explicação anterior. Para o caso dos vídeos, talvez fosse interessante a inserção, no menu, de um tutorial em que tais expressões inéditas fossem apresentadas, bem como a compreensão que se tem delas no contexto da tradução/interpretação e legenda.
- Não tenho elementos para discernir e opinar sobre a resposta à pergunta 27.
- "É necessário muito cuidado ao se fazer uma tradução de português para libras, pois existem especificidades que não são contempladas em uma língua e são na outra. O interprete precisa conhecer muito bem o conteúdo da informação á ser transmitida , a qual pode ser feita por meio de vídeo ."

5.6 PRODUÇÃO DE LEGENDAS EM LS COM AVATARES

Estas questões tratam dos elementos e etapas envolvidas na produção de legendas em vídeo de LS com intérpretes virtuais.

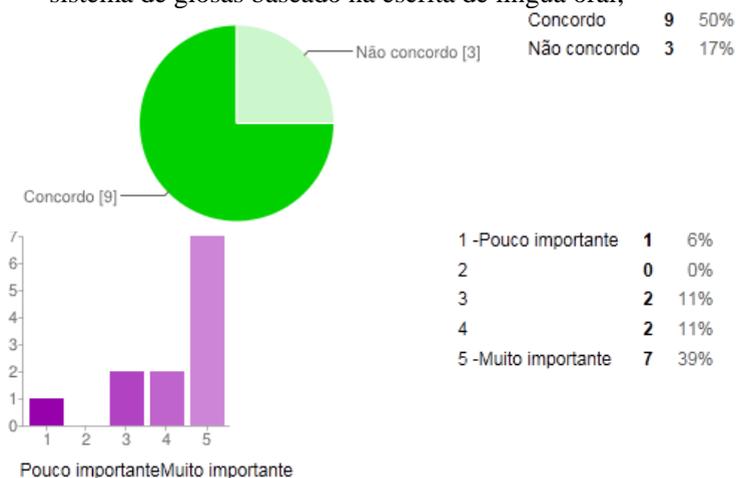
28. As técnicas automatizadas para produção de conteúdo em LS baseiam-se na síntese de sinais por meio da animação de avatares;



(2ª rodada) Diante a grande demanda de material audiovisual a ser traduzido e legendado, e a reduzida disponibilidade de interpretes, os avatares (interpretes virtuais) demonstram ser uma alternativa viável para apoiar a produção automática de conteúdo em LS;



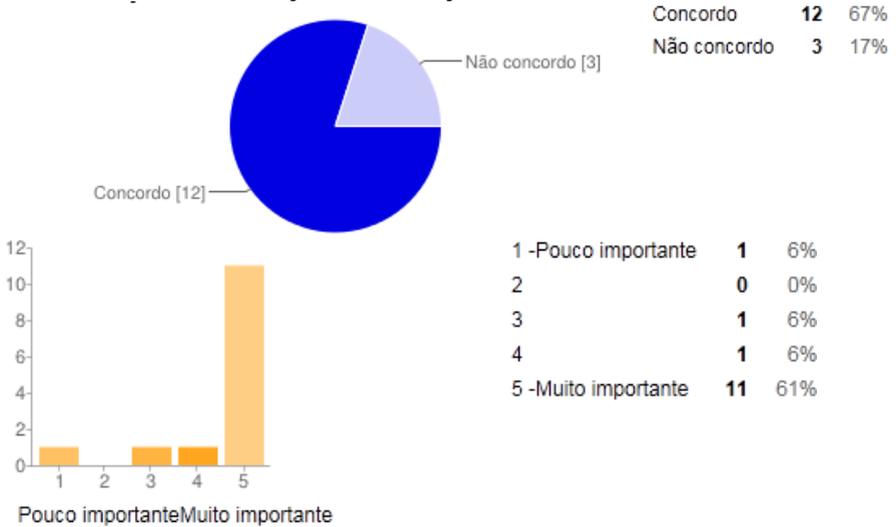
29. A produção de vídeo em LS com avatares depende de um sistema de notação escrita, sendo preferíveis os sistemas voltados às LS ao sistema de glosas baseado na escrita de língua oral;



(2ª rodada) A produção de vídeo em LS com avatares pode utilizar um sistema de notação escrita, sendo preferíveis os sistemas voltados às LS aos sistemas de glosas baseados na escrita de língua oral;



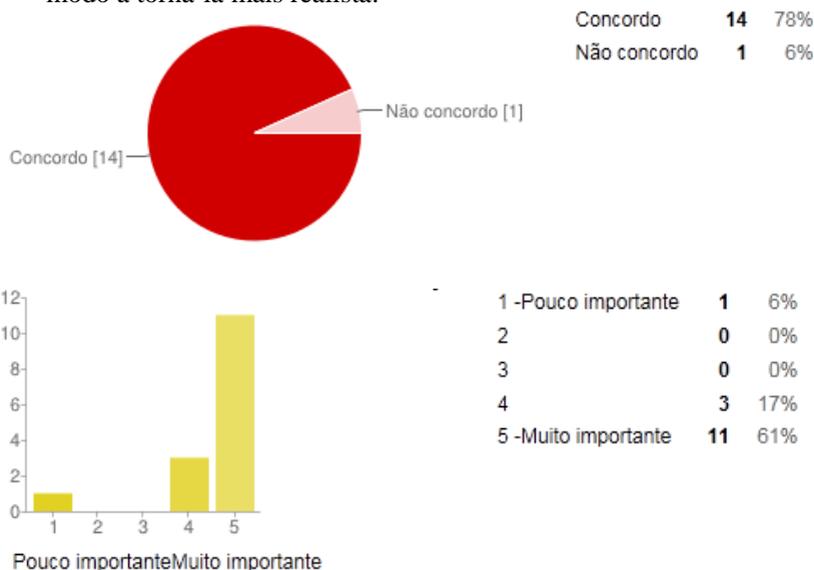
30. Na síntese de sinais por meio de avatares devem ser respeitados os mesmos critérios estabelecidos para o intérprete humano, tais como a vestimenta, iluminação e codificação de vídeo;



(2ª rodada) Os avatares apresentando língua de sinais devem respeitar os mesmos critérios estabelecidos para o intérprete humano, tais como a vestimenta, iluminação e codificação de vídeo;



31. Elementos prosódicos, que indicam o estado emocional ou entonação do intérprete, devem ser aplicados na síntese de sinais de modo a torná-la mais realista.



Comentários sobre Produção de Legendas com Avatares

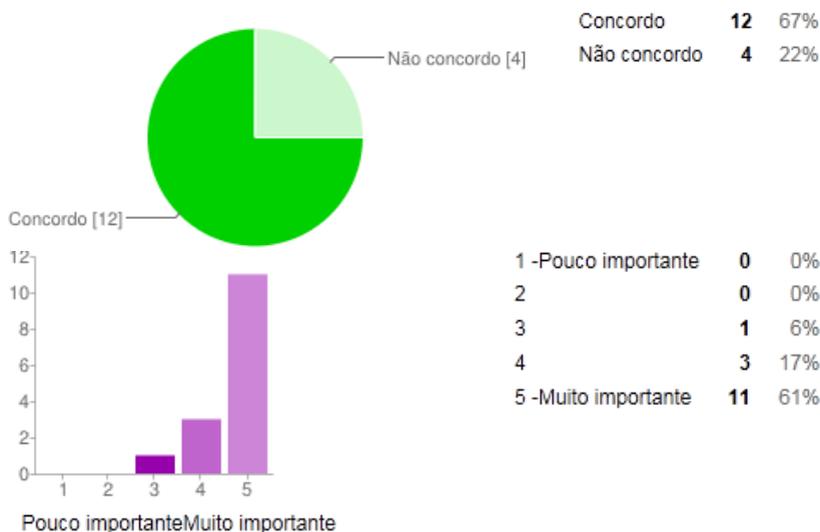
- prefiro considerar ausência de respostas, pois as opções anteriores não caracterizam minha compreensão sobre o tema, pois não considero pouco ou muito importante, mas não conheço tantos recursos que utilizam avatares a ponto de opinar sobre o formato dos mesmos. Tive experiência no uso de um recurso dessa natureza, muito inicial ainda e, embora tenha utilizado o objeto de aprendizagem ao qual me refiro, houve diferença na compreensão de surdos (compreensão maior quando os sinais foram realizados pelo intérprete e menor compreensão quando realizados pelo avatar sobre o mesmo conteúdo). Embora não tenha explorado esses resultados, precisaria de um conjunto maior de conhecimentos nessa área p/ responder com mais propriedade as questões formuladas.
- Não concordo com o uso de avatares.
- prefiro não opinar sobre esse item, pois não conheço essa tecnologia. Se por um lado pode ser atraente para os surdos,

principalmente para as crianças, por outro lado, penso que a figura do interprete pode evitar sistemas de glosa.

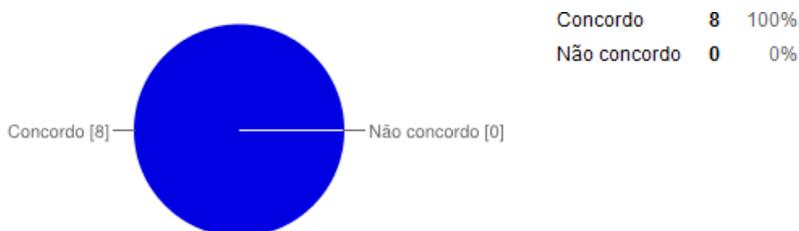
- Falta muitos parâmetros da Libras para a tradução por avatares ser real (2ª rodada).

5.7 PRODUÇÃO DE LEGENDA EM ESCRITA DE SINAIS

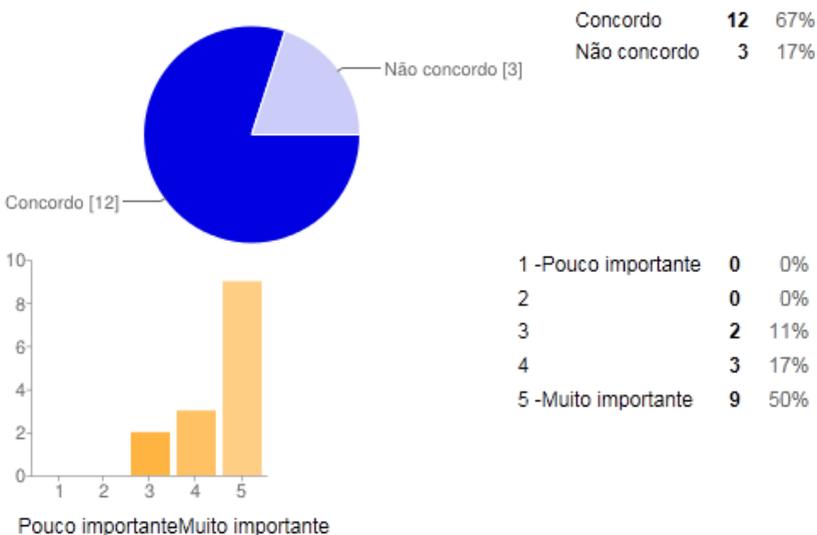
32. As legendas em escrita de sinais podem ser geradas a partir da informação sonora do audiovisual, ou a partir de outras formas de legendas disponíveis;



(2ª rodada) As línguas de sinais são o principal meio de expressão da cultura surda, mas sua forma escrita ainda busca consenso dentro desta comunidade. A escrita de sinais demonstra ser uma importante forma de registro das línguas de sinais, contribuindo para a afirmação da cultura surda. Seu uso nas legendas do audiovisual propicia sua divulgação e estimula seu aprendizado. As legendas em escrita de sinais podem ser geradas a partir das informações do audiovisual, ou a partir de outras formas de legendas disponíveis;

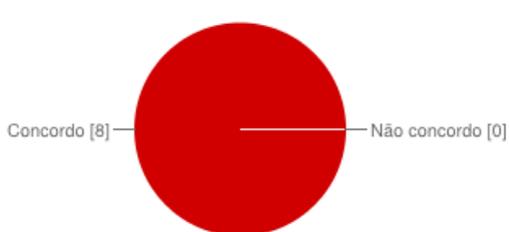


33. Os passos para produção de legenda com escrita de sinais são:
- Análise do conteúdo a ser traduzido;
 - Definição de corpo de linguagem, com inserção de termos no glossário;
 - Tradução português para Libras;
 - Elaboração de texto de sinais;
 - Edição da legenda, com a sincronização e condensação;
 - Revisão da legenda com escrita de sinais;
 - Publicação da legenda.



(2ª rodada) Os passos para produção de legenda com escrita de sinais são:

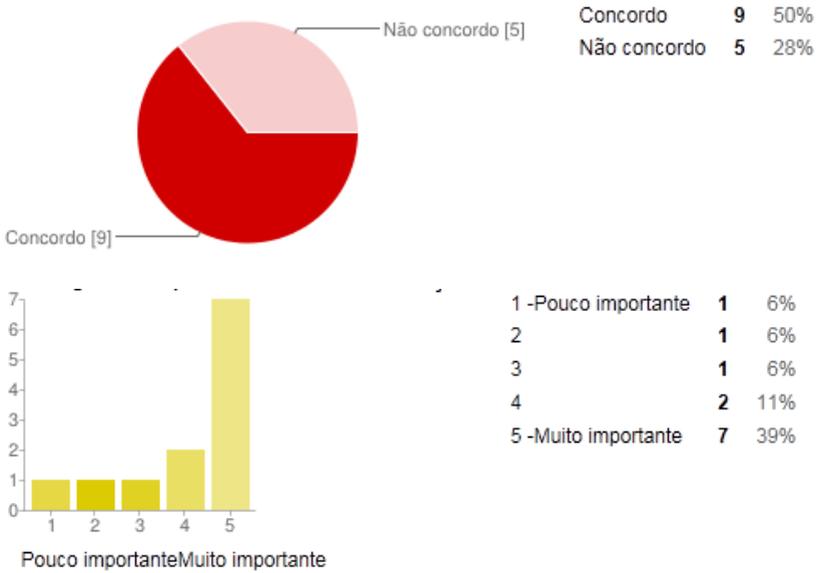
- Análise do conteúdo a ser traduzido;
- Tradução português para Libras, definindo o corpo de linguagem, com inserção de termos inéditos no glossário;
- Elaboração do texto de sinais, podendo utilizar termos do glossário;
- Edição da legenda, com a sincronização e condensação;
- Revisão da legenda com escrita de sinais;
- Publicação da legenda.



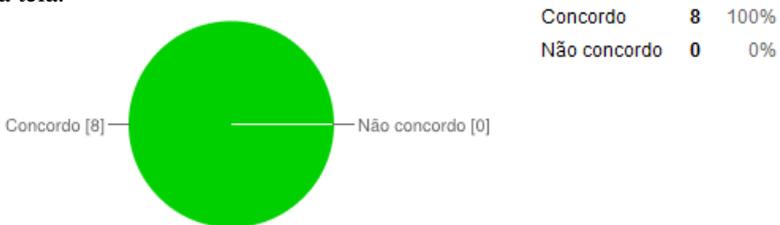
Concordo	8	100%
----------	----------	------

Não concordo	0	0%
--------------	----------	----

34. A apresentação das legendas em escrita de sinais pode ser feita sobre uma área da tela que ocupe um quarto da largura e 100% da altura, e permanece alinhada a direita.



(2ª rodada) É necessário definir parâmetros para a apresentação das legendas em escrita de sinais: tais como o tamanho da escrita, cores, densidade de símbolos exibidos e localização. Por exemplo, esta poderia ser feita sobre uma área da tela que ocupe um quarto da largura e 100% da altura, e permanece alinhada a uma das laterais da tela.



35. Legendas em escrita de sinais oferecem um meio para manifestação da cultura surda, entretanto necessitam de validação quanto a sua efetividade na comunicação.



Comentários sobre desenvolvimento de legendas em escrita de sinais

- Em alguns momentos fiquei em dúvida sobre a opção a marcar. Acho que seria interessante incluir a opção. Não sei/desconheço. Algo do tipo.
- Escrita de sinais tem amplitude e eficácia necessária para uma proposta desse tamanho?
- Há normalização específica para o tamanho e disposição da legenda em LS, pode sofrer variação, então precisa prever isso, mas independentemente de todas as etapas p/ a elaboração da legenda, é fundamental o conhecimento prévio do perfil do público-alvo, pois a maior ou menor compreensão tb está relacionada a esse aspecto.
- Escrita de sinais é mais um recurso de separação entre as línguas e vai contra o bilinguismo. O surdo tem que se concentrar em aprender a ler e escrever bem o português.
- Eu penso que a escrita de sinais pode ser um bom recurso para facilitar a compreensão de surdos, mas eu sou mais favorável à legendas com língua de sinais e legendas em português escrito, pois

a escrita de sinais é mais um fator restritivo para a inserção dos surdos na sociedade.

- Os próprios surdos divergem quanto ao uso da escrita de sinais (2ª rodada).

5.8 VALIDAÇÃO ESTATÍSTICA

A validação estatística utilizou o teste dos sinais baseado na distribuição binomial para amostras pequenas. Este teste busca validar variáveis qualitativas com duas categorias, uma positiva e outra negativa, no caso, cada uma das recomendações apresentadas a partir do modelo.

Foram aceitas as recomendações com $\alpha < 0,01$, na primeira rodada, e com $\alpha < 0,05$ na segunda rodada, conforme quadros do anexo B.

As recomendações não aceitas na primeira rodada foram reestruturadas em função dos comentários e novamente avaliadas. Como resultado, as recomendações 28 e 29, a respeito do uso de avatares e de sistemas de notação de sinais para sua animação, não puderam ser validados no tratamento estatístico.

5.9 CONCLUSÕES SOBRE A CONSULTA AOS ESPECIALISTAS

A partir dos comentários dos avaliadores pode-se perceber a diversidade de visões a respeito de quais artefatos devem apoiar o acesso dos surdos. Há comentários que remetem à cultura de oralização do surdo, ao indicar que deveriam ser usadas apenas legendas em língua oral em programas para adultos. Mas que também respeitam a constituição bilíngue, ao sugerir que programas infantis deveriam ter legendas em LS.

É possível perceber a preocupação dos pareceristas quanto ao uso de legendas em português como meio para que o surdo aprenda o português visando sua inclusão na sociedade ouvinte.

Nessa situação há necessidade de uma solução de compromisso entre a simplificação do texto, de modo a facilitar a leitura, e a manutenção da versão original, de modo a não alienar o surdo de um vocabulário real e mais amplo.

Porém é incompatível o uso exclusivo do texto como meio de acesso, pois neste “momento histórico o uso exclusivo de legenda escrita priva os surdos do acesso à informação”.

A ideia de oferecer legendas com LS em programas infantis enquanto que os programas adultos utilizem mais as legendas orais escritas, está ligada à ideia de incluir o surdo na sociedade ouvinte. Entretanto não considera a possível preferência dos surdos pelas LS mesmo que estes conheçam a escrita do português.

A intenção de disponibilizar legendas em LS na forma escrita é de manter um espaço linguístico das LS, ao mesmo tempo em que beneficia o processo de produção por meio da maior flexibilidade do meio de representação escrito em função do vídeo. Há entretanto falta de comprovação prática para afirmar que esta abordagem de legendagem é funcional, e também sua implementação dependerá do conhecimento desta modalidade de escrita por parte do público alvo.

A falta de comprovação prática em relação à escrita de sinais pode ser atribuída a fatores como o surgimento recente de sistemas de escrita que são efetivamente utilizados pelos surdos, e a ausência de suporte tecnológico para seu uso associado ao audiovisual.

A consulta aos especialistas levou ao aprimoramento do modelo apresentado. As modificações concentram-se na relação entre os papéis desempenhados para a produção de legendas, nos processos de produção de legendas em português escrito, com o uso de avatares e em escrita de sinais.

A figura 92 apresenta a nova versão do modelo de casos de uso decorrente dos comentários dos avaliadores.

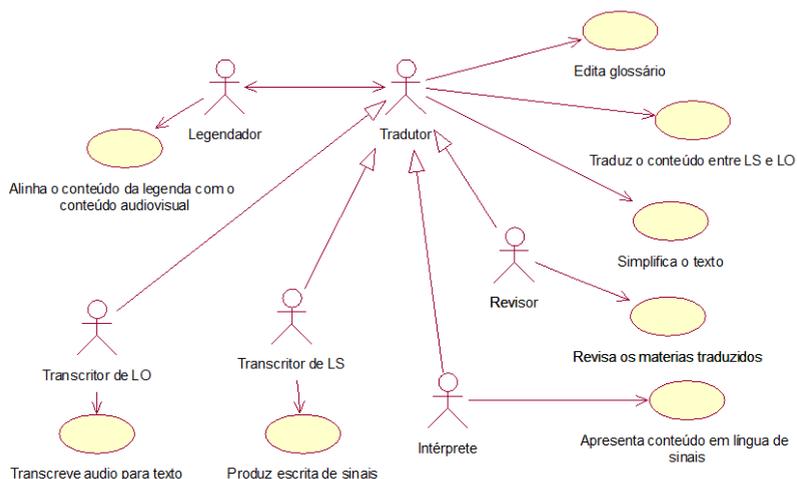


Figura 92: Modelo de casos de uso para produção de artefatos de apoio atualizado.

A principal modificação está no relacionamento entre o papel de legendador e o de tradutor, e na incorporação do papel de um revisor responsável pela validação das legendas produzidas. Este revisor é concebido como um tipo de tradutor.

Estes papéis podem ser desempenhados por uma ou mais pessoas e também podem ser realizados por meio de dispositivos automáticos.

O próximo capítulo descreve a implementação de um sistema de apoio à produção de artefatos para acessibilidade do audiovisual e ilustra a aplicação das recomendações e processos definidos no modelo de referência.

6. IMPLEMENTAÇÃO DE REFERÊNCIA E CENÁRIOS DE USO

Este capítulo apresenta a aplicação do modelo para oferta de audiovisual acessível na forma de cenários de uso e uma implementação técnica de referência.

A implementação de referência aplica as recomendações do modelo no desenvolvimento de uma plataforma que consiste num aplicativo para produção de legendas para surdos e em aplicativos para exibição do conteúdo acessível.

Os cenários de uso ilustram a realização dos processos existentes no modelo. São apresentados três cenários de uso, sendo dois cenários abstratos e um concreto. Os cenários abstratos simulam a aplicação do modelo, e o cenário concreto que demonstra efetivamente a produção de artefatos para acessibilidade a partir da implementação de referência.

O tópico seguinte apresenta a implementação de referência, e os seguintes demonstram os cenários com a aplicação dos processos e artefatos identificados e modelados na pesquisa.

6.1 IMPLEMENTAÇÃO DE REFERÊNCIA

A implementação técnica de referência descreve um protótipo de software de apoio ao desenvolvimento de artefatos para acessibilidade de surdos e sua integração com o audiovisual.

O sistema implementa a estrutura de apoio aos processos apresentados no modelo.

A implementação está estruturada em diferentes módulos que oferecem funcionalidades associadas à produção de legendas para surdos. A figura 93 descreve a arquitetura da implementação, que utiliza recursos descritos anteriormente.

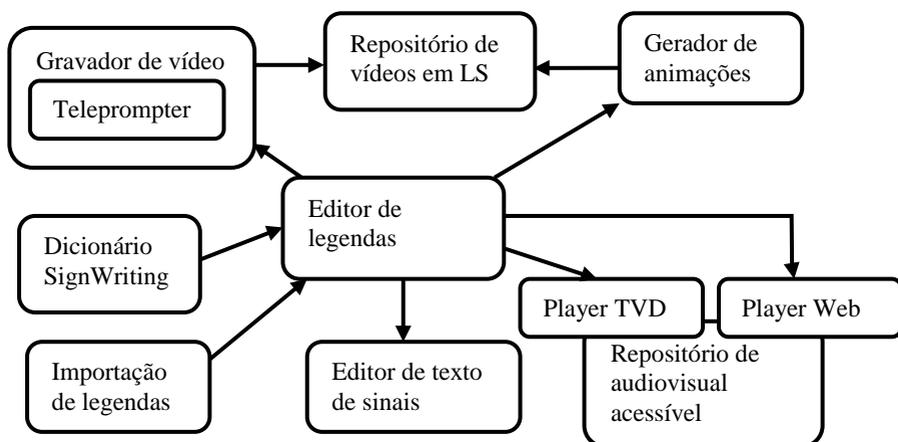


Figura 93: Módulos da arquitetura utilizada na implementação.

O editor de legendas é a interface principal do aplicativo, a partir dele é possível adicionar diferentes recursos de acessibilidade e editá-los. Especificamente é possível adicionar e editar as legendas em texto, em escrita de sinais e em vídeo de LS.

O gravador de vídeo oferece uma plataforma on-line para gravação vídeos em LS a partir de um *teleprompter* ou da exibição de outro vídeo.

Os vídeos gravados são armazenados em um repositório de vídeos de LS, a partir do qual podem ser adicionados ao conteúdo do editor de legendas.

O módulo de importação de legendas permite reutilizar legendas pré-existentes, codificadas tanto no formato da língua oral quanto em escrita de sinais.

O dicionário *SignWriting* é um glossário com a tradução de termos do português para Libras escrito que utiliza a implementação do SignPuddle, e permite a importação de sinais para o editor de legendas.

O editor de texto de sinais é um editor específico para escrita de sinais que permite modificar o conteúdo deste tipo de legendas.

O gerador de animações é responsável por traduzir os dados da escrita de sinais em um vídeo de avatar com LS.

De modo a reutilizar um sistema de gerenciamento de usuários e permissões, a implementação utilizou como base o software Moodle (MOODLE, 2012).

A implementação de referência está disponível no endereço https://github.com/ronniebritto/moodle_moviemasher.

Em relação aos formatos de mídia, o modelo de referência não está focado em apontar diretamente a tecnologias associadas a seus processos, já a implementação de referência pode indicar os possíveis modelos de dados que apoiam os processos identificados no modelo.

O mapeamento das alternativas de modelos de dados relacionados aos artefatos de acessibilidade em conteúdo audiovisual digital é apresentado no quadro 12 a seguir.

Quadro 12: Formatos para apresentação de legendas para surdos

Tipo de legenda	Tipo de dados
Legenda escrita em português	Sub Rip Subtitle Format (SRT)
Legenda de Escrita de Sinais	TSWML
Legenda de avatar sinalizante	MPEG 4 H-Anim
Legenda de vídeo com intérprete em LS	Vídeo com codificação priorizando aspectos faciais e de movimentos manuais (h.264 com AdaptSynch)

Como alternativas para codificação de legendas escritas estão a Timed Text Markup Language (TTML, 2010) e o SubRip Subtitle Format (SRT) . A TTML é uma linguagem de marcação para legendas sincronizadas definida pelo consórcio de padronização web W3C, por sua vez, o SRT é um padrão de fato para o armazenamento deste tipo de legendas. O formato escolhido para legendas em português é o .SRT.

Para as legendas de vídeo com interprete de LS é adotada a codificação h.264 (ITU, 2012) a qual pode ser codificada de acordo com o critério de qualidade sem priorização de conteúdo.

O conteúdo de legendas com avatares pode ser armazenado e transmitido por meio do formato H-Anim como demonstrado na literatura.

Para legendas em LS na forma escrita, foi desenvolvido o formato TSWML (*Timed SignWriting Markup Language*) que é uma extensão da SWML. O formato SWML codifica os símbolos dos caracteres referentes ao sinal que se deseja apresentar. Esta linguagem tem quatro variações: SWML-Brazil, que é o formato SWML original; o SWML-D, utilizado nos dicionários on-line WoordenBoek; o SWML-S, utilizado nos dicionários SignPuddle; e o SWDB, utilizado na base de dados SignWriting (SIGNPUDDLE, 2012).

O formato é estendido com o objetivo de permitir a sincronia temporal com o vídeo principal por meio da adição de atributos de tempo de início e fim para cada conjunto de sinais contidos numa Signbox. Um documento no dialeto SWML-S é apresentado na figura 94, onde pode ser notada a presença dos atributos *begin* e *end* das linhas 4 e 17.

```

1  <?xml version="1.0"?>
2  <!DOCTYPE swml SYSTEM "http://www.signpuddle.com/swml/swml-s.dtd">
3  <tswml dialect="S" version="1.1" lang="sgn" glosslang="">
4  <signbox begin="00:01:29.45" end="00:01:32.89">
5  <sign lane="0">
6  <gloss>sign-5</gloss>
7  <symbol x="86" y="67">03-03-001-01-01-01</symbol>
8  <symbol x="38" y="75">01-03-001-01-01-01</symbol>
9  <symbol x="108" y="108">02-01-001-01-01-01</symbol>
10 </sign>
11 <sign lane="1">
12 <gloss>sign-2</gloss>
13 <symbol x="68" y="66">03-05-001-01-01-01</symbol>
14 <symbol x="42" y="109">01-03-001-01-01-01</symbol>
15 </sign>
16 </signbox>
17 <signbox begin="00:01:33.45" end="00:01:34.89">.....
18 <sign lane="-1">
19 <gloss>sign-3</gloss>
20 <symbol x="67" y="80">03-05-001-01-01-01</symbol>
21 <symbol x="20" y="135">01-04-001-01-01-01</symbol>
22 </sign>
23 </signbox>
24 </tswml>
25

```

Figura 94: Exemplo de código para legenda em escrita de sinais.

Esta modificação constitui o formato *Timed SignWriting Markup Language* (TSWML).

O próximo tópico descreve os módulos que foram implementados a partir do modelo.

6.1.1 Editor de legendas

A principal interface para o usuário baseia-se num editor de vídeo, para o qual duas opções foram analisadas para sua implementação: o uso do software ELAN (ELAN, 2012) e o Moviemasher (MOVIEMASHER, 2012).

Ao analisar o ELAN identificou-se certa dificuldade para a inserção de uma linha de tempo gráfica, que deveria conter a escrita de sinais. Em contato com o desenvolvedor do software o mesmo orientou sobre o processo de realização desta modificação, entretanto esta possibilidade foi evitada devido à falta de integração da plataforma Java com o recurso de webcams.

A necessidade de gravação de vídeo on-line levou à escolha da plataforma Adobe Flash. Concentraram-se os esforços no software Moviemasher. Sua interface pode ser vista na figura 95.



Figura 95: Interface do software para composição de vídeos Moviemasher. Fonte: MovieMasher (2012).

O aplicativo foi modificado de modo a atender os processos identificados no modelo de referência. Foram adicionadas as seguintes funcionalidades para edição de legendas:

- Importação do vídeo principal;
- Importação de texto para *teleprompt*;
- Importação de texto em sinais;
- Importação de legenda em escrita de sinais;
- Importação de legenda em escrita oral;

- Importação de vídeo em LS;
- Gravação de vídeo em LS a partir de *teleprompt*;
- Gravação de vídeo em LS a partir de vídeo;
- Edição de glossário;
- Busca de itens no glossário;
- Geração de animação de avatar a partir de legenda escrita de sinais;
- Geração de animação a partir de legenda em texto oral;
- Geração de texto de sinais a partir de vídeo em LS;
- Exportação do conteúdo para plataformas web e TVD.

A figura 96 apresenta a tela principal do editor de vídeo após sua modificação.

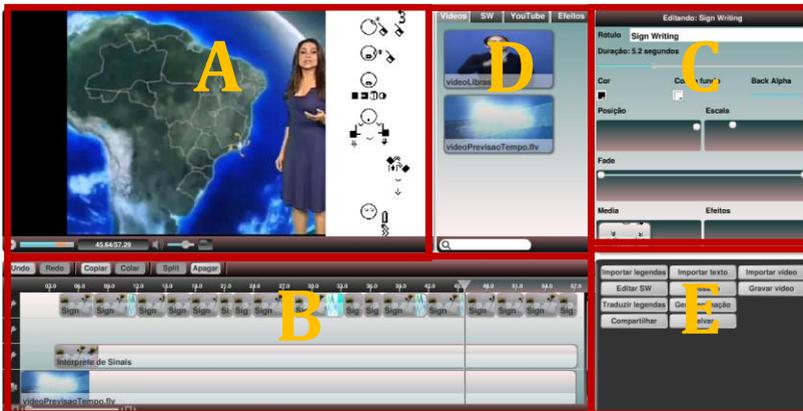


Figura 96: Interface do Moviemasher adaptada à tarefa de tradução/ composição de audiovisual acessível para operação de sistema desenvolvido a partir do modelo de referência proposto.

A interface está estruturada em uma área de player de vídeo (A), linha do tempo para adição de efeitos (B), área de edição de propriedades de efeitos (C), área de seleção de mídia (D), e uma barra de botões de comandos (E).

O player permite a pré-visualização dos efeitos de mídia adicionados.

A área para seleção de mídia esta dividida em SL, SW, Vídeos e Efeitos. SL são vídeo gravados ou importados por meio do editor, SW permite a busca de termos em um glossário com escrita de sinais, a partir de glosas em português. A guia para vídeo possibilita selecionar a

mídia do audiovisual principal. Já a guia de Efeitos permite adicionar os diferentes tipos de legendas à linha do tempo.

A linha de tempo é dividida em camadas, onde a linha inferior corresponde ao conteúdo principal e as linhas superiores representa os artefatos de mídia adicionados de modo a oferecer a acessibilidade. Por exemplo uma linha para legendas textuais com a transcrição literal da fala, outra camada com legenda simplificada, uma para escrita de sinais e outra para vídeo de intérprete.

A área de edição de propriedades de mídia serve para definir parâmetros de exibição, como duração, posicionamento, e efeitos de transição, como *fading* e *blending* dos artefatos presentes na linha do tempo.

Na barra de comandos estão dispostos botões que permitem acionar funções como salvar, gerar animação a partir da escrita de sinais, exportar para NCL, exportar para HTML, importar arquivo de legendas ou texto, além de comandos relacionados à edição da linha do tempo.

O esquema de dados interno do editor de legendas é apresentado no próximo tópico.

6.1.1.1 Esquemas de dados do editor de legendas

O editor de legendas utiliza o esquema de dados definido pelo *Moviemasher*. A estrutura principal do conteúdo produzido pelo editor é o *mash*. Um *mash* é a estrutura que contém todas as informações sobre o audiovisual a ser entregue. Dentro de cada *mash* é possível a definição de *clips*. Cada *clip* possui um efeito definido por um identificador.

Os identificadores disponíveis no contexto do modelo de referência proposto contemplam efeitos para legendas em texto oral, em vídeo de LS e em escrita da LS.

Um trecho de legenda em texto oral pode ser obtido com o trecho de código apresentado na figura 97.

```
<clip type="effect" id="captioneffect" track="3" start="305" longtext="e
elas podem ser mais intensas de Pernambuco ao Rio Grande do Norte"
forecolor="000000" label="Caption" lengthseconds="4.5" padding="3"
backalpha="50" textsize="14" length="45" font="default" bgcolor="FFFFFF"/>
```

Figura 97: Exemplo de dados para legendas em texto oral

A legenda em texto consiste num *clip* com o efeito de *captioneffect* que apresenta o atributo *longtext* a partir do quadro *start*.

Já o efeito de legenda em vídeo de língua de sinais, corresponde a um clip com efeito de *signlanguageeffect*, que apresenta um *clip* de *video*, conforme figura 98.

```
<clip type="effect" id="signlanguageeffect" track="1" start="0" length="543"
label="Sign Language" lengthseconds="54.3">
  <clip type="video" id="afce095da2ead65c23f885e46a6e1df3" audio="0"
trimend="0" volume="" label="Previsão do tempo traduzida para Libras -
Janine" trimstart="0" track="-1" fill="stretch" start="0" length="560"/>
</clip>
```

Figura 98: Exemplo de dados para legendas em LS com vídeo

Por sua vez a legenda em escrita de sinais é representada por um clip to tipo *signboxeffect* que exibe uma imagem definida por um identificador correspondente aos símbolos *SignWriting* que devem ser apresentados, conforme figura 99.

```
<clip type="effect" id="signboxeffect" track="1" start="0" length="34"
label="Sign Language" lengthseconds="3.4">
  <clip type="image" id="0fb33b38c3924f54f521938c3924ff52a16d38d3924fd512"
audio="0" label="Sabado" trimstart="0" track="-1" fill="scale" start="0"
length="34"/>
</clip>
```

Figura 99: Exemplo de dados para legendas em escrita de sinais.

Outro elemento do esquema de dados é o *media*. Cada *media* aponta para um endereço que identifica o recurso a ser apresentado. Cada clip pode ter elementos *media* associados por meio do atributo identificador, sendo possíveis *media* referentes à vídeos, imagens e áudio. A figura 100 apresenta um exemplo de *media* para vídeo.

```
<media group="video" type="video" id="SL7" label="videoLibras.mp4"
duration="55.75"
audio="http://localhost/moodle/file.php/2/moddata/moviemasher/6/2/7.mp4"
url="http://localhost/moodle/file.php/2/moddata/moviemasher/6/2/7.mp4"
icon="http://localhost/moodle/file.php/2/moddata/moviemasher/6/2/7.jpg"
fill="scale"/>
```

Figura 100: Exemplo de dados para mídia de vídeo.

Os tópicos seguintes descrevem a implementação de algumas funcionalidades previstas no protótipo, como a gravação de vídeo de LS, a importação de legendas, a edição de legendas em escrita de sinais e a exportação do conteúdo para exibição na TV digital e web.

6.1.2 Gravador de Vídeo e Teleprompter

Um módulo adicional foi incorporado ao aplicativo de edição de vídeo para permitir a gravação on-line de vídeos em LS. O vídeo em LS pode ser gravado a partir de um *teleprompt* de texto ou então a partir de um vídeo pré-existente.

O aplicativo utiliza o protocolo *Real Time Messaging Protocol* (RMTP) (ADOBE, 2012) que permite o envio e recebimento de fluxos de mídia como vídeo ou áudio. O servidor para este protocolo foi o RED5 (RED5, 2012), por meio de implementação de um método de gravação remoto que é utilizado pelo aplicativo cliente em Flash.

O usuário pode escolher o formato de mídia que deseja gravar, podendo ser o 16:9, 4:3 ou então o formato de janela de LS. Após a gravação o usuário pode revisar o vídeo gravado e gravar um novo caso necessário.

Os vídeos gravados ficam disponíveis na área de seleção de mídia e podem ser associados ao efeito de legenda de LS com intérprete.

6.1.3 Importação de Legendas

Uma funcionalidade incorporada ao aplicativo para atender ao modelo de referência é a possibilidade de importação de legendas pré-existentes. Deste modo é possível importar arquivos com legendas em língua oral ou em LS.

No caso de legendas em português, ao carregar um arquivo .srt o aplicativo converte cada segmento de legenda em um *clip* específico. Especificamente, o seguinte trecho do arquivo SRT

```
35
00:01:29,454 --> 00:01:32,894
de exercitar seus direitos autorais de outros modos, mais
simples.
```

É convertido para o código de representação interna do aplicativo, apresentado a seguir :

```
<clip type="effect" id="captioneffect" track="3"
start="305" longtext="de exercitar seus direitos autorais de
outros modos, mais simples." forecolor="000000" label="Legenda"
lengthseconds="3.4" padding="3" backalpha="50" textsize="14"
length="45" font="default" backcolor="FFFFFF"/>
```

A importação de legendas em escrita de sinais funciona de modo análogo.

O código de representação interna do aplicativo pode ser exportado novamente, gerando o código no formato SRT ou TSWML com as modificações realizadas.

6.1.4 Editor de Legendas de Sinais

Há necessidade de um editor de escrita de sinais online para modificação dos textos em sinais. Para isso o editor *SignText* foi incorporado ao editor de legendas, conforme figura 101. A integração entre o editor de legendas e o editor de texto de sinais utiliza o formato *Binary SignWriting* (SIGNBANK, 2012) para comunicação de dados.

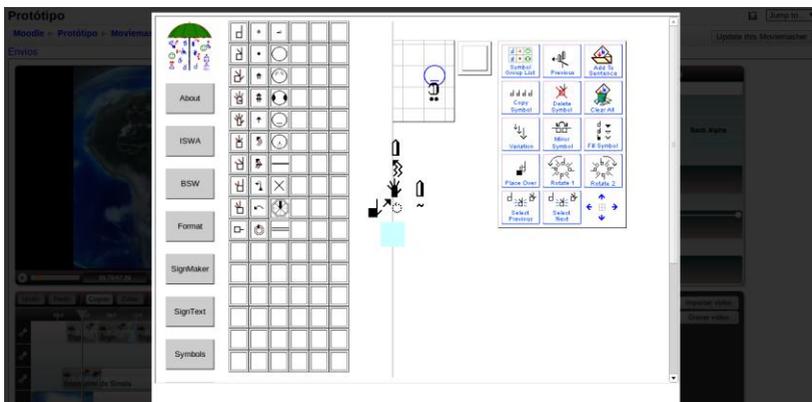


Figura 101: Integração do editor de texto de sinais ao editor de legendas.

Em seu funcionamento, o usuário adiciona a legenda em escrita de sinais à linha do tempo e ativa o editor de texto de sinais para inserir ou editar seu conteúdo.

6.1.5 O dicionário/glossário SignWriting

O módulo referente ao dicionário/glossário SignWriting consiste em uma base de dados que associa termos da língua oral a termos da língua de sinais.

Este dicionário permite que o usuário realize traduções a nível morfológico e também orienta a aplicação dos termos nas legendas, visto que o modelo prevê um glossário que descreva os termos utilizados.

6.1.6 Aplicativo para exibição de legendas

Esta sessão descreve os recursos utilizados na publicação do audiovisual nas plataformas digitais com foco no sistema de TV digital brasileiro, visto que a apresentação do audiovisual na *web* ocorre por meio do player do próprio *Moviemasher*.

Na TV digital, a exibição de arquivos de legenda em texto oral é possível por meio da implementação de um *player* deste tipo de arquivo no respectivo ambiente, como demonstrado por Azevedo (2011). A exibição de legendas em TSWML é realizada com uma modificação deste *player*, especificamente o método que exibia o conteúdo textual passa a exibir imagens a partir da informação contida no arquivo TSWML.

As legendas em língua oral escrita e em LS com intérprete utilizam as recomendações da ABNT. Estes modelos foram apresentados na figura 2 do primeiro capítulo. Já a forma para exibição de legendas em LS na forma de escrita de sinais é apresentado na figura 102, de modo que a área destinada à legenda ocupa um quarto do comprimento da tela e a totalidade da altura, permanecendo alinhada à direita da tela.

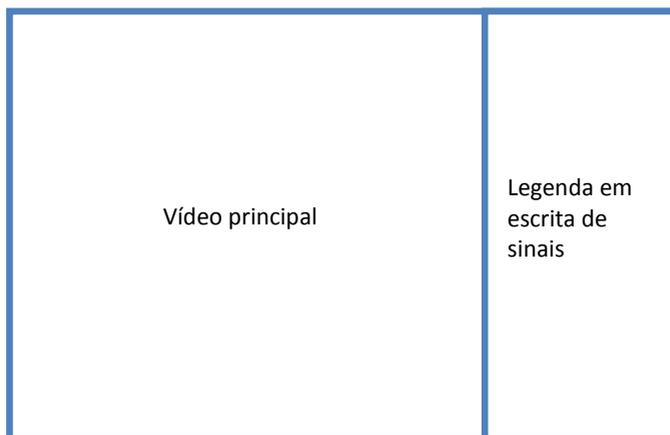


Figura 102: modelo para apresentação de legenda em LS na forma escrita

O acesso ao glossário no caso da *web* é realizado por meio da ativação de um link na interface do usuário. Ao ativar o link para o

glossário, a exibição do vídeo principal é suspensa e retomada após o fechamento do aplicativo de glossário.

Eventos gerados pela interação do usuário com o sistema de TVDI disparam a exibição dos objetos de mídia e também permitem a personalização de sua exibição.

Na implementação foi adotado o modelo de interação descrito no quadro 13. As teclas coloridas exibem os diferentes tipos de legendas e algumas teclas do teclado numérico controlam o posicionamento e tamanho da janela do intérprete de LS.

Quadro 13: Modelo para interação com aplicativo de exibição de legendas.

Botão	Comportamento
Azul	Exibe janela de Libras
Amarelo	Exibe e seleciona legenda em português
Verde	Exibe e seleciona legenda em escrita de sinais
Vermelho	Oculta todos os elementos
Teclado numérico 8	Aumenta o tamanho de exibição
Teclado numérico 2	Diminui o tamanho de exibição
Teclado numérico 6	Desloca a legenda em LS para a direita
Teclado numérico 4	Desloca a legenda em LS para a esquerda

Para a exibição da legenda em LS no formato de vídeo é utilizado um player nativo do sistema de TVDI que exibe arquivos nesse formato. A implementação de referência permite que o usuário ative ou desative a exibição do vídeo, assim como permite seu redimensionamento e reposicionamento na tela.

O esquema de dados do editor/compositor do audiovisual pode ser transformado e traduzido para ser entregue em diferentes dispositivos.

Para a TV digital foi aplicado a linguagem NCL na definição de um aplicativo NCL, o qual contém informações sobre o que tocar, onde, como e quando tocar (Barbosa; Soares, 2008).

A definição de locais de exibição das mídias é feito por meio de regiões. A figura 103 apresenta um exemplo de código que define as regiões onde serão exibidas as legendas em um aplicativo da plataforma de TV digital.

```

<regionBase>
  <region id="rgVideoPrincipal" left="0%" top="0%" width="80%" height="80%" zIndex="1" />
  <region id="rgVideoLibras" left="0%" top="50%" width="25%" height="50%" zIndex="3" />
  <region id="rgLegendasTexto" left="0%" top="0%" width="100%" height="59%" zIndex="2" />
  <region id="rgLegendasSW" left="75%" top="0%" width="25%" height="100%" zIndex="2" />
  <region id="rgLua" left="75%" top="0%" width="25%" height="100%" zIndex="3" />
</regionBase>

```

Figura 103: Código de definição das regiões de exibição das legendas na TV digital

A tela é dividida em regiões, havendo uma região dedicada ao fluxo principal do audiovisual, e outros que definem as dimensões correspondentes a cada tipo de legenda.

A definição sobre o que tocar é feita por meio do elemento *media*, conforme figura 104, que apresenta as mídias do vídeo principal, composta pelo fluxo principal, a mídia para a legenda, composta por um vídeo no formato *mpeg4*, e as legendas textuais, que são exibidas por meio de *scripts* na linguagem Lua (PUC, 2012).

```

<media id="videoPrincipal" src="sbtvd-ts://0.ts" descriptor="dVideoPrincipal">
  <property name="bounds"/>
</media>
<media id="legendaVideoLibras" src="media/videoLibras.mpeg" descriptor="dVideoLibras">
  <property name="bounds"/>
</media>
<media id="legendaTexto" src="srtPlayer/mediaplayer_srt.lua" descriptor="dLegendasTexto">
  <property name="path"/>
  <property name="fps"/>
  <property name="color"/>
</media>
<media id="legendaTextoSinais" src="tswmlPlayer/mediaplayer_tswml.lua" descriptor="dLegendasSW">
  <property name="bounds"/>
</media>
<media id="luaController" type="application/x-ginga-NCLua" src="controller.lua" descriptor="dLua">
  <property name="startVideoPrincipal"/>
  <property name="startLibras"/>
  <property name="startLegendaTexto"/>
  <property name="startLegendaTextoSinais"/>
  <property name="stopAll"/>
</media>

```

Figura 104: Código para definição de mídias para acessibilidade de surdos na TV digital.

A mídia de com *id="luaController"* é um *script* NCL/Lua responsável por controlar a interação com o usuário, capturando eventos da interface e acionando ou desativando a exibição das legendas.

Em complemento, elementos descritores associam as mídias às regiões de exibição e permitem definir os parâmetros de exibição.

O aplicativo de exibição de legendas para TV digital foi executado no OpenGinga (OPENGINGA, 2012), que oferece uma implementação do middleware de TV digital brasileira executável em computadores PC.

Os próximos tópicos descrevem cenários de uso que ilustram a aplicação da implementação de referência.

6.2 CENÁRIOS DE USO

Cada cenário de uso está estruturado em um momento de produção de conteúdo e em outro de fruição do conteúdo. A fase de produção visa apresentar os recursos para o desenvolvimento de conteúdo acessível, e a fase de fruição ilustra como estes conteúdos são efetivamente acessados.

Os cenários de uso estão divididos em situações onde o conteúdo é apresentado na web e cenários com conteúdo para TVD.

Os cenários de uso são:

- a) Cenário de vídeo educativo na web: Legendagem com escrita de sinais de um vídeo educativo na web;
- b) Cenário de telenovela na TVD: legendagem com texto em português e vídeo de LS;
- c) Cenário previsão do tempo na web: Legendagem com intérprete real de LS e escrita de sinais de uma previsão do tempo;
- d) Cenário de transcrição de vídeo de LS para legenda em escrita de sinais.

6.2.1 Cenário de uso com vídeo educativo na web

Este cenário de uso visa explorar os meios de automação dos processos de produção de artefatos para acessibilidade.

O objeto de aplicação dos métodos e técnicas é um vídeo produzido por um projeto educacional que oferece recursos de acessibilidade.

6.2.1.1 Produção

A trilha de áudio é extraído do arquivo do vídeo e passa por um processo de reconhecimento automático de áudio, que gera uma legenda com a transcrição textual na língua oral com anotações sobre aspectos prosódicos de cada frase.

Cada trecho da legenda textual em língua oral é traduzida para LS, gerando uma legenda com escrita de sinais e outra com um intérprete virtual.

As legendas são incorporadas ao vídeo principal, sendo então revisadas. Os trechos com problemas ou erros de tradução e

interpretação são corrigidos e realimentam o sistema de tradução automática.

Para os erros na transcrição do áudio é utilizado o editor de legendas de texto e para problemas de tradução do texto da língua oral para o avatar com LS é utilizado um editor de animação para alterar os dados de movimentos.

6.2.1.2 Fruição

O usuário inicia a exibição do vídeo na web e ativa a legenda textual, porém não se sente confortável com o texto na língua oral.

Ele clica no ícone de ativação da legenda em LS e passa a assistir a legenda com interprete virtual.

Aos 48 segundos do vídeo ele não entende um sinal realizado pelo avatar, nesse momento o vídeo é parado e é feita uma consulta ao glossário de LS utilizando como parâmetro de busca os componentes do sinal que foi apresentado pelo avatar.

A figura 105 descreve os elementos do modelo de referência utilizados nesse cenário de uso.

Utilizou-se o reconhecimento automático de áudio o qual gerou um texto na língua oral, que foi sincronizado e simplificado de modo a gerar a legenda em texto oral.

Um processo de tradução automática, com o uso de um glossário/dicionário, gera uma legenda em escrita de sinais e outra com intérprete virtual.

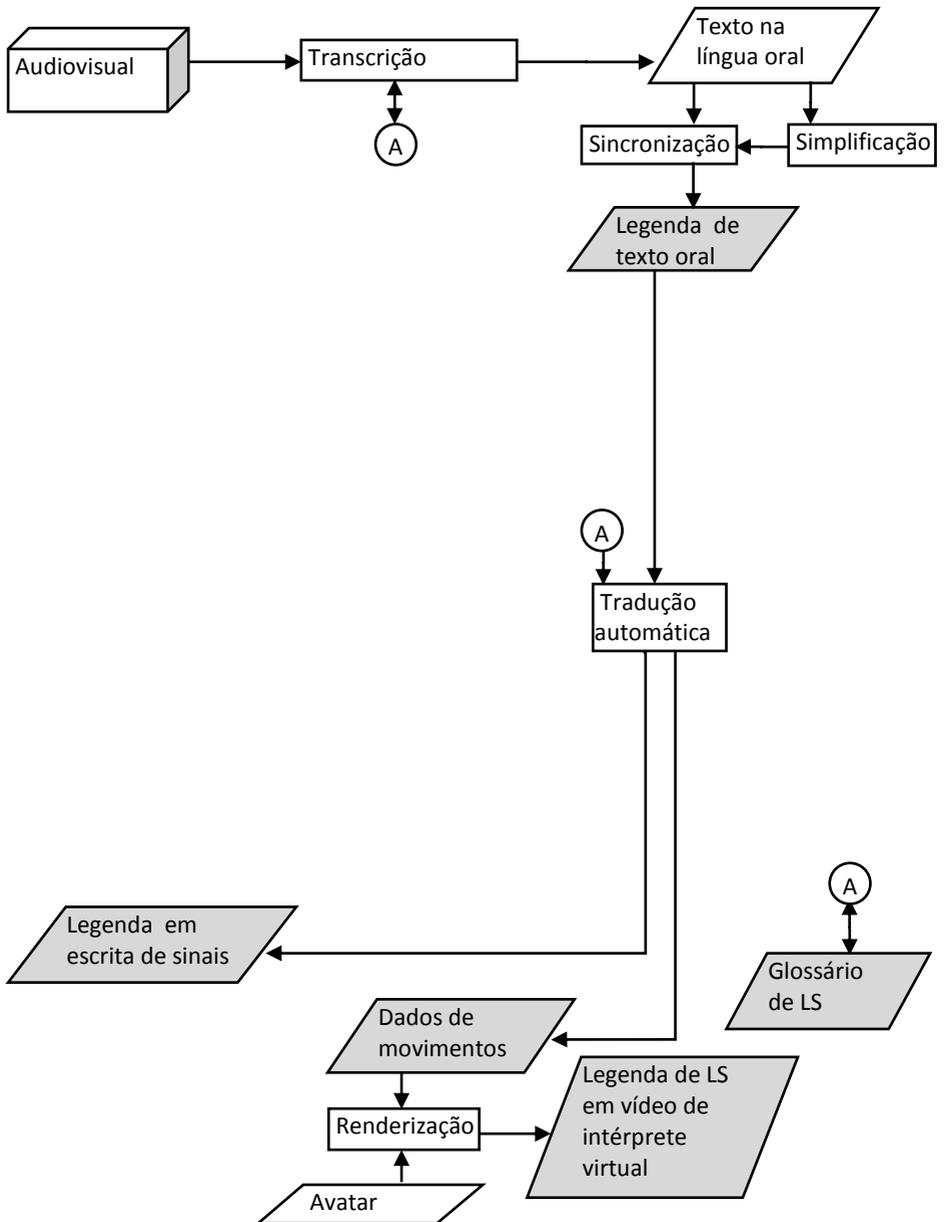


Figura 105: Elementos de modelo de referência utilizados no cenário de uso.

6.2.2 Cenário de Telenovela na TVD

Este cenário descreve a produção de legendas para uma telenovela exibida na plataforma de TVD.

6.2.2.1 *Produção*

O capítulo da telenovela tem duração de 50 minutos, e já foi pré-produzida e gravada em alta definição de imagem.

O texto com o script dos atores está disponível para a equipe responsável por tornar o conteúdo acessível.

Este texto escrito é importado no editor de legendas e sincronizado com as falas de cada personagem em cada cena. Trechos que superam o limite de 120 caracteres por minuto passam por um processo de condensação.

Anotações sobre efeitos sonoros das cenas são adicionados.

Após a produção da legenda em texto com a transcrição das falas das personagens, as legendas passam por um processo de simplificação adicional, que visa reduzir ao máximo a quantidade de caracteres apresentados.

Este processo da origem a dois níveis de legendas textuais: uma transcrição textual com eventual limitação de caracteres por minuto, e a legenda condensada.

Após a criação das legendas multinível, estas são encaminhadas para uma máquina de tradução automática, que gera um texto em LS apresentado em vídeo com intérprete virtual.

As legendas produzidas são empacotadas e enviadas via ISDB-T.

6.2.2.2 *Fruição*

O espectador surdo começa a assistir a telenovela e habilita a exibição de legendas em texto do português. Entretanto considera muito rápida a velocidade com que são exibidas e ativa a legenda com nível maior de condensação.

Durante a leitura da legenda o surdo não entende a palavra ‘variação’. Ele pausa a exibição do vídeo e consulta um glossário bilíngue, que apresenta uma descrição do termo em português e em LS.

Após um certo tempo assistindo a novela e encontrando muitas palavras desconhecidas no português ele percebe que aumentou seu vocabulário tanto em LS quanto em português.

A figura 106 descreve os elementos do modelo de referência utilizados nesse cenário de uso. A partir de um texto em língua oral preexistente ocorreu a sincronização e simplificação para manter a quantidade de caracteres adequada. Em seguida um processo de tradução automática gera um vídeo com intérprete virtual de LS que é usado como legenda.

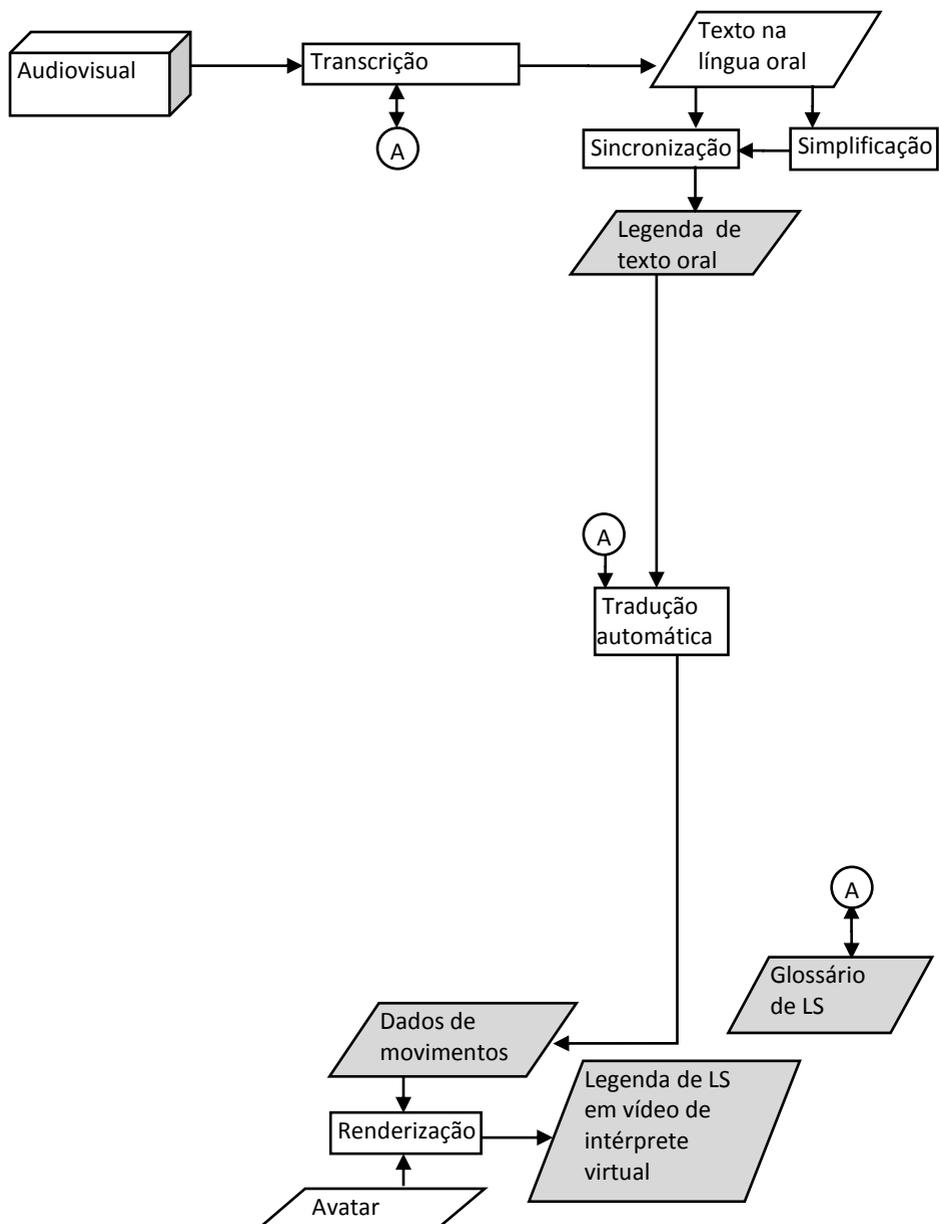


Figura 106: Elementos de modelo de referência utilizados no cenário de uso.

6.2.3 Cenário de Previsão do Tempo na Web

O objetivo deste cenário é explorar os recursos de vídeo com TILS e da escrita de sinais em um cenário de publicação de um vídeo na web, que foi anteriormente exibido ao vivo na televisão.

6.2.3.1 Produção

Uma emissora apresenta a previsão do tempo para o país no jornal do meio dia, a apresentadora da previsão aponta para regiões do mapa enquanto fala. A duração do vídeo é de um minuto.

Após ir ao ar, o conteúdo é encaminhado para produção de legendas, de modo a ser publicado no ambiente web.

A intérprete de LS assiste ao vídeo da previsão e identifica os sinais que devem ser apresentados, consultando um glossário para tirar dúvidas sobre os sinais para os estados brasileiros.

Em seguida parte para um estúdio onde faz a gravação do vídeo usando a técnica de *Serial Relay*, filmando o conteúdo em LS sem o auxílio de *teleprompter*. O vídeo gravado é codificado priorizando-se a região facial.

O vídeo em LS é encaminhado para transcrição por um especialista em escrita de sinais e cria um texto com o aplicativo SW-Edit, apresentado na figura 107. A transcrição de LS é posteriormente importada no editor de legendas e sincronizada com vídeo em LS.

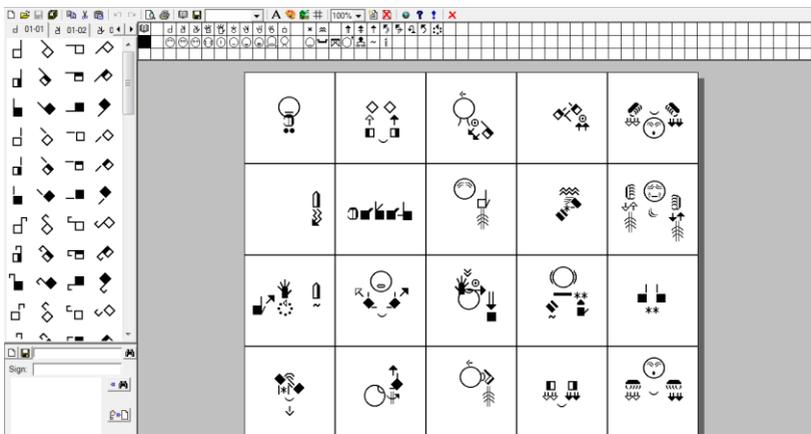


Figura 107: Editor de texto SW-Edit.

O vídeo com LS e as transcrições sincronizadas com o vídeo principal são revisados pelos responsáveis pelas traduções.

Constata-se que a legenda em escrita de sinais apresenta muitos caracteres que são exibidos simultaneamente. Ela então é editada de modo a exibir de quatro a seis símbolos simultâneos.

Após a condensação as legendas são aprovadas e encaminhadas para publicação pelo setor responsável por conteúdo web da emissora.

O setor de conteúdo publica o vídeo principal em conjunto com as legendas por meio de um tocador especialmente projetado para tocar legendas em escrita de sinais.

A produção do material ocorreu pela tradução por meio de intérprete do áudio original, para a produção de legenda com vídeo em LS, a partir do qual foi gerado o texto em escrita de sinais, que pode ser vista no anexo C.

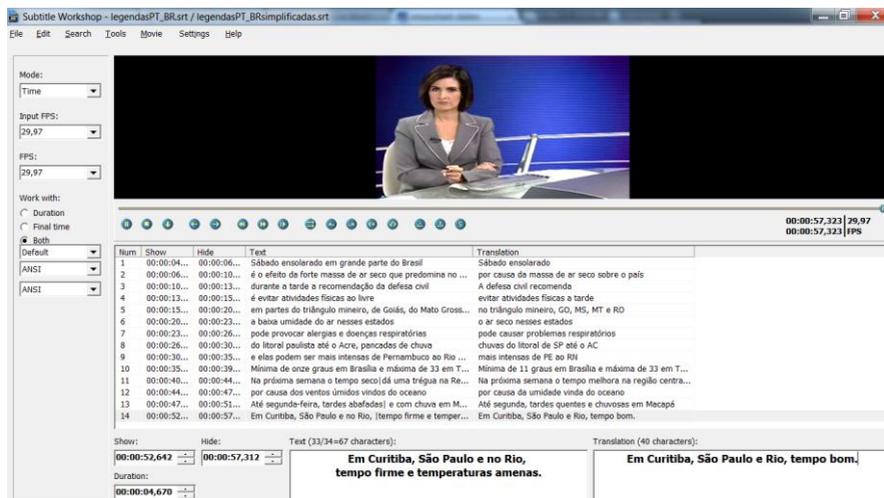


Figura 108: interface do aplicativo de produção de legendas *Subtitle Workshop*.

O aplicativo *Subtitle Workshop* (URUSOFT, 2012), apresentado na figura 108, serviu de apoio para a produção de legendas multinível. A simplificação do texto ocorreu de modo a reduzir informações redundantes e omitir informações desnecessárias, sendo que algumas palavras, referentes aos nomes de estados, foram substituídas por suas siglas. O quadro 14 apresenta a correspondência entre o conteúdo das duas legendas.

Quadro 14: Exemplo de diferentes níveis de condensação de legendas

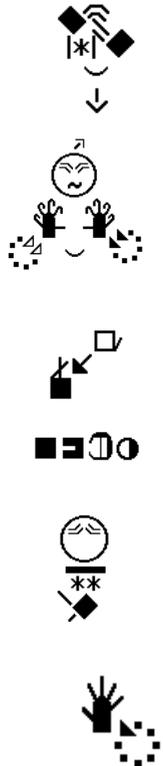
Tempo	Legenda com transcrição	Legenda simplificada
00:00:04,126 --> 00:00:06,885	Sábado ensolarado em grande parte do Brasil	Sábado ensolarado
00:00:06,886 --> 00:00:10,851	é o efeito da forte massa de ar seco que predomina no país	por causa da massa de ar seco sobre o país
00:00:10,852 --> 00:00:13,752	durante a tarde a recomendação da defesa	A defesa civil recomenda

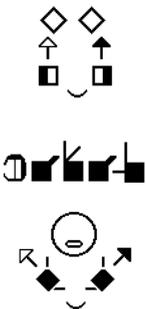
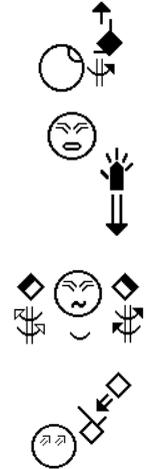
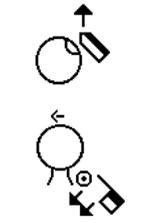
	civil	
00:00:13,753 --> 00:00:15,963	é evitar atividades físicas ao livre	evitar atividades físicas a tarde
00:00:15,964 --> 00:00:20,854	em partes do triângulo mineiro, de Goiás, do Mato Grosso do Sul, do Mato Grosso e de Rondônia	no triângulo mineiro, GO, MS, MT e RO
00:00:20,855 --> 00:00:23,475	a baixa umidade do ar nesses estados	o ar seco nesses estados
00:00:23,691 --> 00:00:26,624	pode provocar alergias e doenças respiratórias	pode causar problemas respiratórios
00:00:26,625 --> 00:00:30,065	do litoral paulista até o Acre, pancadas de chuva	chuvas do litoral de SP até o AC
00:00:30,650 --> 00:00:35,109	e elas podem ser mais intensas de Pernambuco ao Rio Grande do Norte	mais intensas de PE ao RN
00:00:35,110 --> 00:00:39,360	Mínima de onze graus em Brasília e máxima de 33 em Teresina.	Mínima de 11 graus em Brasília e máxima de 33 em Teresina
00:00:40,131 --> 00:00:44,943	Na próxima semana o tempo seco dá uma trégua na Região central do país,	Na próxima semana o tempo melhora na região central do país,
00:00:44,944 --> 00:00:47,560	por causa dos ventos úmidos vindos do oceano	por causa da umidade vinda do oceano
00:00:47,561 --> 00:00:51,471	Até segunda-feira, tardes abafadas e com chuva em Macapá	Até segunda, tardes quentes e chuvosas em Macapá
00:00:52,642 --> 00:00:57,312	Em Curitiba, São Paulo e no Rio, tempo firme e temperaturas amenas.	Em Curitiba, São Paulo e Rio, tempo bom.

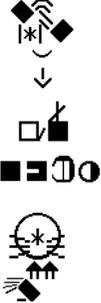
Araujo (2008) recomenda que para cada 1s de fala, sejam gerados 14 caracteres. O texto possui 793 caracteres ditos em 58 segundos, o que corresponde a uma taxa de 13,7 caracteres por segundo, dentro da recomendação. O quadro 17 apresenta o texto traduzido, tanto em língua oral quanto em LS. As glosas foram geradas pela tradução fornecida por dois leitores de escrita de sinais.

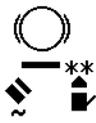
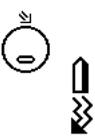
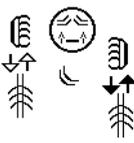
Quadro 15: Correspondência entre conteúdo da transcrição em língua oral e em LS

Tempo	Duração	Transcrição original	Glosas	Transcrição em escrita de sinais
-------	---------	----------------------	--------	----------------------------------

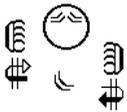
<p>00:05</p>	<p>2s</p>	<p>Sábado ensolarado em grande parte do Brasil</p>	<p>Sábado; Brasil; Sol/Todo Brasil</p>	
<p>00:07</p>	<p>3s</p>	<p>é o efeito da forte massa de ar seco que predomina no país</p>	<p>Por causa; massa de ar; Ar seco; Tem; Área / região</p>	

00:10	3s	durante a tarde a recomendação da defesa civil	Defesa; Civil Informar/Avisar	
00:13	3s	é evitar atividades físicas ao livre	Evitar Tarde Atividade física Sol	
00:16	3s	em partes do triângulo mineiro, de Goiás, do Mato Grosso do Sul,	Principal Minas Gerais Goiás	

			Mato Grosso do Sul	
00:18	2s	do Mato Grosso e de Rondônia	Mato Grosso Rondônia	
00:21	3s	a baixa umidade do ar nesses estados	Por causa Ar seco Respirar ar seco	
00:24	3s	pode provocar alergias e doenças respiratórias	Problema Doença Pulmão/Costas	

			Pode	 
00:27	4s	do litoral paulista até o Acre, pancadas de chuva	São Paulo Brasil Até/geral (desenha) Acre Chuva Chuva sim/ chuva não	     
00:30	5s	e elas podem ser mais intensas de Pernambuco ao Rio Grande do	Também Chuva	

		<p>Norte</p>	<p>Forte</p> <p>Pernambuco/Re cife</p> <p>até o Rio Grande do Norte</p>	
<p>00:35</p>	<p>5s</p>	<p>Mínima de onze graus em Brasília e máxima de 33 em Teresina.</p>	<p>Mínima/ Graus</p> <p>Onze</p> <p>Brasília</p> <p>Máxima /</p>	

			Macapá	
00:52	6s	Em Curitiba, São Paulo e no Rio, tempo firme e temperaturas amenas.	Curitiba São Paulo Rio Tempo Bom	    

6.2.3.2 *Fruição*

Um espectador surdo que não conseguiu assistir a previsão do tempo ao vivo, pois não tinha acesso às legendas, acessa o site da emissora e navega até o conteúdo que deseja assistir.

Ele ativa o recurso de legendas em vídeo de LS e clica no botão de tocar. Aos 37 segundos ele não reconhece o sinal correspondente a palavra “Brasília”. Ele pausa a exibição do vídeo e consulta o glossário de LS, que apresenta a descrição do sinal, com sua forma escrita e uma aplicação da palavra em um contexto.

O usuário retoma a exibição do vídeo e repete a consulta ao glossário para outros sinais que não conhecia.

Após assistir à previsão do tempo, resolve assisti-la uma segunda vez, agora ativando a exibição simultânea de legenda em vídeo e legenda em escrita de sinais, conforme figura 109.



Figura 109: Exemplo de fruição de conteúdo com legenda em escrita de sinais na plataforma de TV digital.

A figura 110 apresenta os elementos do modelo de referência utilizados neste cenário de uso.

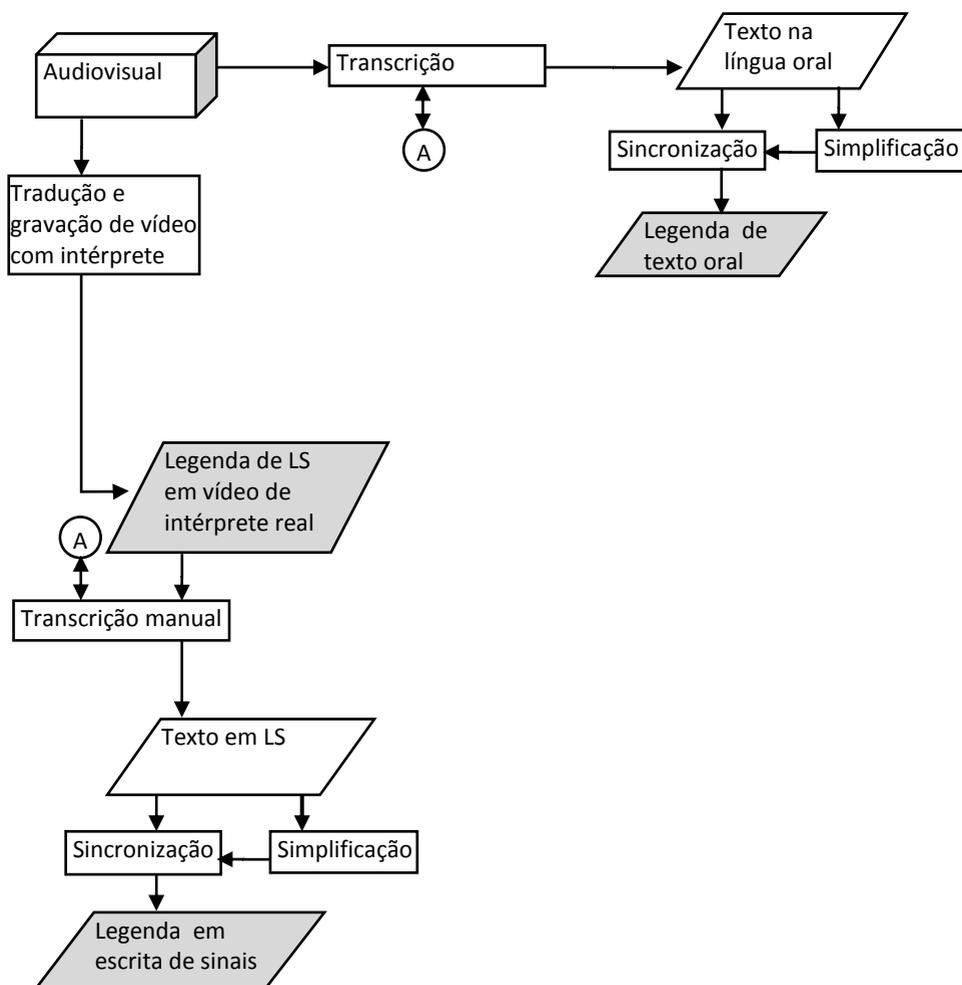


Figura 110: Elementos do modelo utilizados no cenário de uso.

Este cenário de uso adotou os caminhos pela tradução manual. Houve a transcrição manual dos diálogos com a respectiva simplificação e sincronização, o que gerou dois níveis de legendas em texto oral.

Para a legenda em LS foi utilizada a abordagem de tradução e gravação manual de vídeo de LS, o qual após sua revisão gerou uma legenda em LS que foi transcrita manualmente e originou um texto em

LS. Este texto passou por um processo de fragmentação e sincronização com a legendas em LS e gerou a legenda em escrita de sinais.

6.2.4 Cenário de transcrição de vídeo em LS

Este cenário de uso surgiu durante a interação com os pesquisadores de escrita de sinais Valerie Sutton, inventora do sistema SignWriting, Stephen Slevinski e Adam Frost, programadores de software de apoio ao uso de escrita de sinais em ambientes digitais.

Estes foram convidados a conhecer e utilizar o protótipo gerado pela implementação de referência e referendar a possibilidade de uso de escrita de sinais na legendagem.

Adam Frost utilizou o sistema e enviou um arquivo de vídeo em LS com legenda em língua inglesa. Ele transcreveu o conteúdo em escrita de sinais utilizando a implementação de referência, a qual pode ser vista na figura 111.



Figura 111: Transcrição bilíngue criada por Adam Frost.

A respeito do uso de escrita de sinais na legendagem Adam, afirmou que “Sim, eu penso que é muito viável realizar traduções de uma língua falada para uma língua de sinais usando SignWriting. É basicamente o mesmo princípio do que eu fiz só que o vídeo principal não teria um sinalizante. ;-)”.

Este cenário de uso descreveu o uso de escrita de sinais nas legendas e utilizou os processos previstos no modelo apresentados na figura 112.

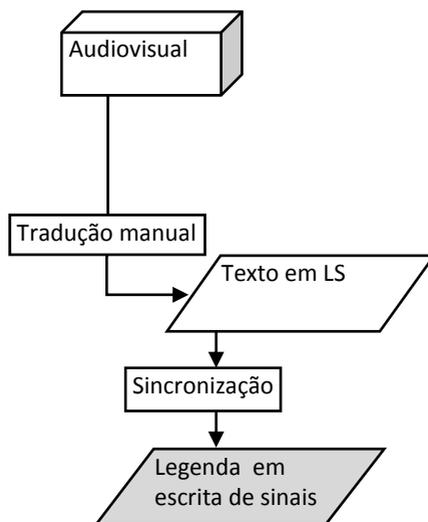


Figura 112: Elementos do modelo utilizados no cenário de uso.

O próximo tópico apresenta as conclusões a respeito do processo para a implementação de referência e dos cenários de uso.

6.3 CONSIDERAÇÕES SOBRE A IMPLEMENTAÇÃO DE REFERÊNCIA E CENÁRIOS DE USO

O modelo para produção de artefatos que promovem a acessibilidade dos surdos foi aplicado em uma implementação de referência, o qual teve sua aplicação descrita nos cenários de uso.

A implementação utilizou recursos preexistentes, como o editor de vídeo *Moviemasher*, o editor de texto de sinais *SignText* e o dicionário *SignPuddle*.

Algumas funcionalidades previstas no modelo não foram implementadas, mas seu acesso está previsto na interface do aplicativo de produção de legendas.

Os primeiros cenários de uso contextualizam a aplicação de trabalhos relacionados à produção automatizada de legendas em português e em Libras.

O terceiro cenário demonstrou em maiores detalhes o desenvolvimento dos processos manuais apresentados no modelo de referência e gerou os artefatos de legenda em texto oral, em texto de sinais e em vídeo com intérprete de LS.

A exportação dos artefatos produzidos para as plataformas digitais consiste na conversão do formato de dados do editor para os formatos específicos de cada plataforma.

A padronização dos formatos de dados para os conteúdos das legendas deve considerar o objetivo de comunicação e o meio utilizado. No caso, os formatos apresentados no tópico correspondente se relacionam ao texto em língua oral, no formato .srt; ao texto em língua de sinais, formato TSWML; ao vídeo, no formato h.264; e aos dados de animação de avatares, que podem usar o formato H-Anim.

O próximo capítulo apresenta as conclusões do trabalho e recomendações para desenvolvimentos futuros.

7. CONCLUSÕES

A abordagem metodológica utilizada na pesquisa permitiu alcançar os objetivos propostos por meio da busca sistemática da bibliografia, que resultou numa base de documentos que descrevem pesquisas em áreas associadas aos surdos e o audiovisual. Estas iniciativas foram enumeradas e caracterizadas, que associadas à experiência do pesquisador em ambiente de pesquisa e ensino voltado ao público surdo, forneceram fundamentos teóricos e práticos para a proposição de um modelo de referência. Este modelo foi descrito por meio de casos de uso, classes e atividades, e foi avaliado, validado e aprimorado com o método Delphi. Após a validação, os processos e artefatos identificados foram implementados em um protótipo que foi aplicado em diferentes cenários de uso, os quais ilustraram a aplicação dos conceitos desenvolvidos no trabalho.

Em relação aos objetivos específicos, a caracterização dos diferentes perfis de surdos, realizado no segundo capítulo, pode ser feita diante de fatores físico-biológicos, relacionados ao tipo e momento do surgimento da surdez, e fatores linguístico-culturais, que representam o ambiente social em que ocorre o desenvolvimento do sujeito surdo. Estes fatores têm ampla variação e possibilidade de combinação, desta forma percebe-se que não é possível determinar um único tipo de surdo.

A oferta de acessibilidade deve considerar que a escolha dos tipos de recursos a serem usados para um acesso efetivo deve ser feita pelo próprio surdo. Pode existir o surdo ensurdecido, que tem preferência exclusiva pela legenda em língua oral escrita, e em outro extremo o surdo que não possui referências à língua oral e prefere usar a LS para sua comunicação.

Assim é possível identificar duas opções principais em relação às possibilidades para comunicação: por língua de sinais ou por meio da língua oral.

Diante as diferentes visões a respeito da surdez e de preferências dos surdos adotou-se um enfoque que oferece recursos para acesso ao audiovisual tanto por meio da língua oral quanto por meio de LS. As possibilidades para desenvolvimento de artefatos de apoio fundamentaram-se na ideia de que o surdo pode conhecer tanto a língua portuguesa, para interação com a sociedade dos ouvintes, quanto a língua de sinais, como meio de comunicação efetivo no contexto da cultura surda. Este enfoque é inerente à acessibilidade do audiovisual visto que este é produzido em uma língua oral e o acesso dos surdos é ampliado quando seu conteúdo é traduzido para uma língua de sinais.

Os artefatos que permitem o acesso de surdos ao audiovisual foram identificados e caracterizados no terceiro capítulo. A legenda em texto oral escrito, a legenda em vídeo de LS, e a legenda em escrita de sinais são os artefatos de maior representação. Em complemento estão os glossários bilíngues e os avatares para apresentação das LS.

A comunicação de legendas via texto consome menos dados e é mais flexível quanto às possibilidades de edição do material traduzido, porém é menos ergonômica para alguns surdos.

A disponibilidade de um glossário com as descrições e traduções dos termos utilizados nas legendas amplia o acesso ao conteúdo apresentado no audiovisual.

O vídeo é o modo mais natural para comunicação por línguas de sinais em ambientes digitais, diante sua natureza visual e gestual. Esta modalidade deve considerar os aspectos de qualidade de codificação de vídeo necessários para comunicação por LS.

A comunicação via vídeo em línguas de sinais, é mais agradável e pode ser feita com o uso de uma câmera e um editor de vídeo, porém demanda mais dados para uma apresentação adequada das LS e possui restrições quanto ao anonimato e edição do conteúdo.

O uso da escrita de sinais como artefato de promoção da acessibilidade depende de sua confirmação como uma forma eficiente de apresentar conteúdo em LS associado ao audiovisual. Contribuem para isso a educação bilíngue e a alfabetização de surdos com uso de escrita de sinais. Além disto, a oferta de recursos tecnológicos para a apresentação dos sistemas de escrita de sinais fomenta o uso desta modalidade de comunicação.

Os avatares estão associados às técnicas de tradução automática, embora também possam ser utilizados em técnicas manuais, apresentando também maior escalabilidade para aplicação em diferentes plataformas de entrega. Em sua aplicação devem ser considerados aspectos referentes à qualidade da animação, evitando-se comportamentos robóticos por meio da aplicação de elementos para variação da entonação dos sinais.

A entrega de conteúdo em LS por meio do envio de dados de movimentos do intérprete virtual e a posterior renderização no receptor do usuário é vantajosa do ponto de vista da possibilidade de configuração de aspectos da exibição do intérprete virtual e da economia da banda de transmissão do canal de TV ou do servidor web, entretanto esta abordagem exige maior capacidade de processamento dos receptores.

No cenário de TV digital o uso de vídeos e avatares pode ser comprometido pela incapacidade do hardware do receptor em exibir o conteúdo. Desde a impossibilidade de exibição opcional e sobreposta de dois vídeos sincronizados, até a demora para converter os parâmetros de animação em um vídeo com o avatar, que devem ser exibidos em tempo real.

No caso do envio de conteúdo pré-renderizado em vídeo, este exige maior demanda de dados, porém menos processamento. Um ponto importante está na sua sincronia com o conteúdo audiovisual e na qualidade gráfica de sua exibição final.

A integração dos processos e mídias existentes para produção destes artefatos em um modelo de referência foi realizada e apresentada por meio de diagramas e quadros apresentados no quarto capítulo.

O modelo de referência proposto permite identificar quais métodos e técnicas podem ser utilizados em sistemas para produção de artefatos de apoio à acessibilidade dos surdos ao conteúdo audiovisual digital. Para representação do modelo foram modelados os casos de uso, classes de entidades e atividades de processos voltados à acessibilidade do surdo ao audiovisual.

Os papéis apresentados nos casos de uso identificam as responsabilidades dos atores envolvidos na produção de recursos para acesso ao audiovisual. Entre os papéis, o tradutor tem grande importância nos processos para elaboração do material que permitirá o acesso do surdo. Evidenciou-se também o acúmulo de funções que podem ser feitas por um ou mais atores, como o transcritor de LO ou LS, que também realizam conversões inter e intralinguísticas. Em complemento ao tradutor, o papel do intérprete é realizar a apresentação dos sinais traduzidos.

Em relação às classes de artefatos para acesso os tipos principais são as legendas e os glossários. As legendas podem ser subclassificadas em tipos voltados à língua oral e às LS.

Classes de entidades relacionadas à produção de LS utilizando a animação de avatares descrevem os tipos de técnicas que podem ser adotadas.

Foram descritas as atividades associadas aos processos de produção de legendas escritas, em vídeo de LS e por meio de avatares.

As possibilidades de processos de tradução foram mapeadas e relacionadas às iniciativas de pesquisa identificadas na revisão bibliográfica.

A adequação do tipo de legenda a diferentes públicos surdos, sua adequação ao tempo de leitura de diferentes usuários, o uso de sistemas

de escrita de LS e a sincronização da legendagem constituem requisitos para potencializar o acesso dos surdos ao audiovisual.

O arranjo dos processos de produção de legendas avaliado por especialistas permite validar as alternativas identificadas para a produção destes materiais. As recomendações derivadas do modelo foram validadas e apresentadas no quinto capítulo por meio de questionários enviados a especialistas relacionados ao tema da pesquisa, que forneceram comentários que aprimoraram o modelo proposto.

O desenvolvimento de uma implementação de referência é demonstrado no sexto capítulo, que ilustra a aplicação de elementos do modelo em um editor de legendas on-line e ferramentas de apoio.

O protótipo gerado pela implementação de referência reutilizou recursos de software como um editor de vídeo on-line, um editor de texto de sinais e um glossário de LS.

Arquivos de vídeo de LS podem ser importados no editor de legendas e editados em relação à linha de tempo do conteúdo. A posição e escala do vídeo na tela principal também podem ser configurados em tempo de edição, e podem ser personalizados pelo usuário na exibição.

Por meio do editor de legendas é possível a realização de atividades identificadas no modelo, como a importação de legendas escritas, a edição de legenda em vídeo de LS e a publicação dos artefatos de apoio.

A extensão do SWML num formato para associação do conteúdo em texto de sinais com o tempo do audiovisual foi definido para permitir a criação de legendas nesse formato.

A implementação de referência também apresentou um aplicativo de exibição de legendas que explora as possibilidades oferecidas pela TV digital. O modelo de interação adotado buscou demonstrar a implementação de recomendações como a possibilidade de escolha do tipo e características das legendas.

A simulação do tocador de legendas no OpenGinga (OPENGINGA, 2012) demonstra a viabilidade de oferta efetiva deste recurso para o sistema brasileiro de TV digital.

Os recursos para a tradução automática não puderam ser amplamente adotados devido à dificuldade de acesso aos aplicativos que implementam as iniciativas encontradas na bibliografia e também pela complexidade técnica em compatibilizar diferentes plataformas, como foi o caso da tentativa de integração do PoliLibras ao protótipo.

Em resumo, os elementos metodológicos e tecnológicos que apoiam a produção e oferta de artefatos para apoio ao audiovisual

puderam ser demonstrados por meio da prototipagem orientada pelo modelo de referência.

Cenários de uso com produção e fruição de conteúdos com acessibilidade aos surdos foram apresentados no sexto capítulo e demonstram a aplicação de elementos do modelo de referência.

Nos cenários de uso buscou-se explorar as possibilidades de alternativas de processos identificados no modelo para produção de artefatos para acessibilidade.

O primeiro cenário descreveu a aplicação de alternativas para a tradução automática, por meio do reconhecimento de áudio para geração de legendas em texto oral, a tradução automática para a geração do vídeo de LS com o uso de avatares e de conteúdo escrito em LS com legenda em texto de sinais.

O segundo cenário demonstrou a possibilidade de reuso do script ou outro material em texto para a produção das legendas em texto oral e sua tradução automática para LS com o uso de avatares.

O terceiro cenário ilustrou de forma concreta o uso do protótipo gerado pela implementação de referência na produção de legenda em LS com intérprete real, em língua oral escrita e em escrita de sinais.

A tradução e interpretação do conteúdo para vídeo em LS foi feita por um tradutor e intérprete de LS e demandou de sua parte grande esforço para realizar a tradução em sincronia com o conteúdo do audiovisual. Este vídeo em LS serviu de base para a produção da legenda em escrita de sinais, que consistiu na transcrição dos sinais realizados pelo intérprete no vídeo.

O quarto cenário de uso surgiu diante a interação com pesquisadores de escrita de sinais e demonstrou a aplicação do modelo e implementação de referência na legendagem em escrita de sinais.

7.1 PESQUISAS FUTURAS

Podem ser identificadas algumas perspectivas para desenvolvimentos futuros a partir dos resultados obtidos com a pesquisa.

Um das possibilidades é a definição de uma plataforma de software colaborativa, orientada pelo modelo de referência, que possibilite a atuação conjunta de tradutores, intérpretes e legendadores para produção de material para acessibilidade de audiovisual.

A tradução de conteúdo para outra língua não é uma tarefa simples, pois demanda conhecimento da cultura de origem e da cultura de destino da tradução. Em geral a tradução automática está longe de ser ideal entre línguas orais, e a situação não é diferente quando se passa de um código oral para um código gestual. O uso de recursos como memória de tradução e a formação de dicionários compartilhados permitirá otimizar este aspecto do desenvolvimento de artefatos para acesso dos surdos ao audiovisual.

No modelo proposto, é utilizado o termo simplificação ou condensação. Deve ser ressaltado que esta ação se refere principalmente à diminuição da quantidade de caracteres apresentado ao usuário surdo e não busca de forma alguma formar um espectador menos informado e sim tornar as legendas mais ergonômicas à leitura de alguns usuários.

Considera-se que desenvolvimentos e validações de aplicativos baseados no modelo proposto possam abrir a possibilidade de estudos mais aprofundados sobre variados aspectos que atualmente não são tão fácil devido à falta das mídias adequadas. Como por exemplo, aspectos de design dos elementos visuais referentes às legendas nos audiovisuais, como tamanho, tipo e cor de fontes tipográficas para textos em LO e de caracteres gráficos das escritas de sinais. Estudos sobre tipo de movimentos de legendas que melhor se adéquam ao usuário também podem acontecer. São necessárias pesquisas também sobre as variações do posicionamento, tamanho e tipo da janela da LS em relação ao tipo de usuário e do conteúdo apresentado.

Uma das grandes limitações na representação das LS é sua dependência do sistema de escrita da língua oral, por exemplo o uso de glosas para identificar os termos que devem ser apresentados na interpretação. Este problema pode ser superado com estudos sobre o uso efetivo dos sistemas de escritas de sinais, que mantém a natureza gestual da comunicação e evitam as perdas decorrentes da tradução interlinguística/semiótica.

Uma abordagem para geração de LS baseada na escrita de sinais pode permitir reduzir os custos de produção das legendas nas línguas de sinais, ao mesmo tempo que respeita a preferência do surdo pelo espaço conceitual delimitado pelo uso de uma língua gestual ao invés da língua oral.

O uso de escrita de sinais nas legendas oferece um ambiente para a padronização nacional dos sinais da Libras, como indicado por AVELAR (2010). Esta padronização está ligada ao uso de um glossário que ofereça acesso universal para a produção de legendas e centralize o cadastro e gerenciamento dos termos utilizados.

Apesar da padronização, a variação regional pode ser contemplada por meio de dicionários que contemplem também definições de sinais regionais.

Há necessidade de padronização tanto dos sinais na sua forma gestual quanto em sua forma escrita, por exemplo a forma de escrita nas legendas quando se utiliza a datilologia, visto por exemplo a diferença de disposição dos símbolos nos termos referentes à “Teresina”, dispostos na vertical, e “ar seco”, dispostos na horizontal, conforme figura 113.

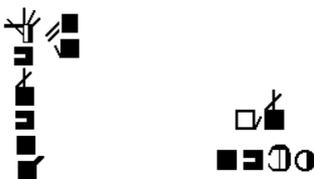


Figura 113: Exemplo de diferença na disposição dos símbolos manuais quando usada a datilologia.

O controle da sincronia de mídias no ambiente de acesso é um fator crítico para a compreensão da legenda e do audiovisual. Configura-se relevante a validação prática de um modelo para interação na plataforma de TV digital que permita a escolha dos níveis de legendas textuais, o redimensionamento da janela de LS e o redimensionamento da área reservada ao conteúdo principal.

Efeitos sonoros do audiovisual podem ser reconhecidos automaticamente, como o barulho de um motor ou o padrão de uma música, e os ícones que os representam podem ser adicionados a uma

linha de efeitos correspondente, a qual poderá gerar os elementos indicativos destes efeitos sonoros para os surdos.

Já os efeitos prosódicos, como entonação e aspectos emotivos, podem também ser reconhecidos e codificados automaticamente. Para esta finalidade há de se estabelecer um vocabulário para representação textual de tais aspectos.

Entre os recursos apresentados, a possibilidade de apresentação de vídeo em LS em conjunto com a legenda em escrita de sinais pode permitir seu aprendizado pelo surdo. Do mesmo modo, a oferta de legendas em português escrito associadas a um glossário bilíngue potencializa o aprendizado da língua oral pelo surdo.

Ao final pode se afirmar que o audiovisual bilíngue, ou multilíngue, configura-se não mais como uma barreira na interação com a sociedade, mas sim uma janela que serve de ferramenta para o aprendizado e interação linguístico-cultural por parte dos surdos com o mundo dos ouvintes, além de possibilitar um espaço para a prática das LS.

7.2 CONSIDERAÇÕES FINAIS

O modelo de referência apresentado descreveu os artefatos que permitem a acessibilidade dos surdos ao audiovisual, e relaciona os processos para a produção destes artefatos. Espera-se que um modelo comum permita o reuso, integração e aprimoramento de iniciativas existentes e a identificação de outras que ainda demandam desenvolvimento.

As restrições relativas às plataformas de entrega tendem a ser superadas pela evolução tecnológica, além disso, a convergência digital exigirá a adoção de padrões interoperáveis para conteúdo audiovisual e também para os recursos de acessibilidade associados. Deste modo um modelo de referência atua como guia para desenvolvimento e integração de métodos e técnicas que geram artefatos de apoio ao acesso do surdo ao audiovisual.

Há de se notar que a produção audiovisual é predominantemente baseada na cultura oral e posteriormente oferecido ao surdo por meio do desenvolvimento de artefatos de apoio ao seu acesso. Porém o contrário também seria interessante, onde o ouvinte é apresentado ao conteúdo audiovisual originalmente sinalizado e inacessível a quem desconhece as LS, mas que torna-se acessível com a dublagem oral ou legendagem escrita.

Ressalta-se o caráter inédito desta tese quando foi apresentado um modelo de referência para integrar formas diferenciadas de legendas adequadas ao tipo de usuário surdo integrando-as aos audiovisuais, prevendo a reutilização de processos, métodos e técnicas para a produção de artefatos que promovam a acessibilidade de surdos aos conteúdos audiovisuais em plataformas digitais.

REFERÊNCIAS

- ABNT. Associação Brasileira de Normas Técnicas. **NBR 15290 – Acessibilidade em comunicação na televisão**. 2005. 10 p.
- ABNT. Associação Brasileira de Normas Técnicas. **NBR 15606 - Codificação de dados e especificações de transmissão para radiodifusão digital**. 2007. 101 p.
- ABNT. Associação Brasileira de Normas Técnicas. **NBR 15604 – Televisão digital terrestre - Receptores**. 2008. 68 p.
- ADAMO-VILLANI, N.; WILBUR, R.B.; ECCARIUS, P.A.; ABE-HARRIS, L.. Effects of character geometric model on perception of sign language animation. In: 2nd International Conference in Visualization, 2009, Barcelona. **Proceedings of the 2nd International Conference in Visualization**. Barcelona, 2009, p. 72-75.
- ADAMO-VILLANI, N.; HAYWARD, K.; LESTINA, J., WILBUR, R.. Effective animation of sign language with prosodic elements for annotation of digital educational content. In: ACM SIGGRAPH 2010 Talks. **Proceedings of SIGGRAPH '10**, Nova Iorque, 2010, p.39.
- ADAMO-VILLANI, N.; BENI, G. Sign language subtitling. In: ACM SIGGRAPH 2005 Educators Program. **Proceedings of SIGGRAPH '05**, Nova Iorque, 2005, p. 1-5.
- ADOBE. Real-Time Messaging Protocol specification. Versão 1.0. Disponível em < <http://www.adobe.com/devnet/rtmp.html>>. Acesso em 15 fev 2012.
- AGRAFIOTIS, D.; CANAGARAJAH, N.; BULL, D.R.; KYLE, J.; SEERS, H.; DYE, M.. A video coding system for sign language communication at low bit rates. In: International Conference on Image Processing. **Proceedings of ICIP'04**. Singapore, 2004, pp. 441-444.
- AHMED, A. S.; SEONG, D. S. K.. SignWriting on mobile phones for the deaf. In: ACM International Conference Proceeding Series. **Proceedings of the 3rd international conference on Mobile technology, applications & systems**. Bangkok, 2006.

ALBRES, N. A. Tradução em Língua Brasileira de Sinais de texto informativo televisivo: reflexões sobre o processo. **Revista Eletrônica de Linguística Domínios de Linguagem**, Uberlândia, v.4, n.1, p. 131-150, set. 2010.

ALLEN, J.M.; ASSELIN, P.K.; FOULDS, R.. American sign language finger spelling recognition system. In: Northeast Conference in Bioengineering. **Proceedings of the Northeast Conference in Bioengineering**, 2003, pp. 285-286.

ALVARENGA , A.; CARVALHO, P. S.; ESCÁRIA, S.C.E.. **DELPHI – Métodos e Aplicações. Ed. Departamento de Prospectiva e Planeamento e Relações Internacionais. Divisão de Informação e Comunicação. Lisboa. 2007.** Disponível em <<http://www.dpp.pt/Lists/Pesquisa%20Avanada/Attachments/3022/Delphi.pdf>>. Acesso em 01 fev 2011.

AMARAL, W. M. ; DE MARTINO, J. M. ; ANGARE, L. M. G. Sign Language 3D Virtual Agent. In: 5th International Multi-Conference on Society, Cybernetics and Informatics. **Proceedings of IMSCI'11**, v.1., Orlando, Julho, 2011, p. 93-97.

AMORIN, M.L.C.; ASSAD R.; LÓSCIO, B.F. FERRAZ, F.S.; MEIRA, S. RybenáTV: Solução para acessibilidade de surdos para TVDigital. In: Simpósio Brasileiro de Sistemas Multimídia e Web. **Anais do Simpósio Brasileiro de Sistemas Multimídia e Web**, Belo Horizonte, 2010. p.243- 248.

ARAÚJO, V.L.S.. **Por um Modelo de Legendagem para Surdos no Brasil.** Tradução & Comunicação: Revista Brasileira de Tradutores, n. 17, 2008.

ARAÚJO,T.M.U.; FILHO, G.L.S; TAVARES, T.A.. An architecture to generate automatic Brazilian sign language legends into Digital Television Systems. In: EuroITV 2009 Networked Television. **Proceedings of EuroITV**, 2009.

ARCOVERDE, R.D.L.. **Digital technologies: A new interactive space of social in the written production of the deaf** [Tecnologias digitais: Novo espaço interativo na produção escrita dos surdos] (2006) Cadernos

CEDES, v.26, n. 69, pp. 251-267.

AVELAR, T.F.. **A Questão da Padronização Linguística de Sinais nos atores-Tradutores Surdos do Curso de Letras-Libras da UFSC: estudo descritivo e lexicográfico do sinal “cultura”**. Dissertação (Mestrado em Estudos da Tradução). Universidade Federal de Santa Catarina. Florianópolis, 2010.

AZEVEDO, R.. Legenda Interativa: **A Shared Culture**. Disponível em <<http://clube.ncl.org.br/node/131>>. Acesso em 15 out 2011.

BARBOSA, S.D.J.; SOARES, L.F.G.. TV digital interativa no Brasil se faz com Ginga: Fundamentos, Padrões, Autoria Declarativa e Usabilidade. Em T. Kowatowski; K. Breitman (orgs.). **Atualizações em Informática**. Rio de Janeiro, RJ: Editora PUC-Rio, 2008. pp 105-174.

BESKOW J. **Talking Heads - Models and Applications for Multimodal Speech Synthesis**. 63 f. Tese (Doutorado) . Institutionen för talöverföring och musikakustik; 2003. Trita-TMH, 2003:7.

BAECKER, R. **The Web of Knowledge Media Design Highlights of a speech given by Professor Ron Baecker Director**, Knowledge Media Design Institute On 23 January 1997 at the OISE Auditorium Toronto, Canada. Disponível em <http://www.dgp.toronto.edu/people/RMB/kmdi_talk.pdf>. Acesso em 15 out 2009.

BAIN, K., BASSON, S., WALD, M.. Speech recognition in university classrooms. In: Fifth International ACM SIGCAPH Conference on Assistive Technologies. **Proceedings of the Fifth International ACM SIGCAPH Conference on Assistive Technologies**, ACM Press, 2002, p. 192-196.

BALDASSARI , S.; CERESO, E. ; SERON, F. J. Maxine: A platform for embodied animated agents. **Computers&Graphics**, v. 32, p. 430-437, 2008.

BALDESSAR, M.J.; GIGLIO, K. **O papel dos sistemas digitais de televisão na economia do conhecimento**. FAMECOS, v. 17, v.1, 2010.

BARROS, MARIÂNGELA ESTELITA. **Escrita das Línguas de Sinais: proposta teórica e verificação prática**. Tese. (Doutorado em Linguística) – Universidade Federal de Santa Catarina, Florianópolis, 2008

BARTH, C.. **Construção da leitura/escrita em Língua de Sinais de crianças surdas em ambientes digitais**. Dissertação (Mestrado em Educação). Universidade Federal do Rio Grande de Sul. Programa de Pós-graduação em Educação. Porto Alegre. 2008. 141p.

BBC. **British Broadcasting Corporation: IPlayer Signed TV**. Disponível em <<http://www.bbc.co.uk/iplayer/tv/categories/signed>>. Acesso em 15 jan 2011.

BRASIL . LEI Nº 10.436, DE 24 DE ABRIL DE 2002. Dispõe sobre a Língua Brasileira de Sinais - Libras e dá outras providências.

BRITO, R.F.; PEREIRA, A.T.C. A model to support sign language content development for digital television. In: Multimedia Signal Processing, 2009. **Proceedings of IEEE International Workshop on Multimedia Signal Processing, MMSP '09**. Rio de Janeiro, p.1-6, out. 2009.

BRITO, L. F. **Por uma gramática de língua de sinais**. Rio de Janeiro: Tempo Brasileiro, UFRJ, Departamento de linguística e filosofia, 1995.

BUEHLER, P.; EVERINGHAM, M.; ZISSERMAN, A.. Learning sign language by watching TV (using weakly aligned subtitles). In: IEEE CVPR, Miami, junho, 2009.

BUNGEROTH, J.; STEIN, D. ; DREUW, P.; ZAHEDI, M. ; NEY, H.. **A German Sign Language Corpus of the Domain Weather Report**. In LREC 06, The fifth international conference on Language Resources and Evaluation, Genoa, Italy, May 2006.

BURNHAM, D., LEIGH, G., NOBLE, W., JONES, C., TYLER, M., GREBENNIKOV, L., VARLEY, A. **Parameters in television captioning for deaf and hard-of-hearing adults: Effects of caption rate versus text reduction on comprehension**. Journal of Deaf Studies and Deaf Education, 13 (3), pp. 391-404. (2008)

BUTTUSSI, F., CHITTARO, L., COPPO, M. **Using Web3D technologies for visualization and search of signs in an international sign language dictionary.** Proceedings of the Twelfth International Conference on 3D Web Technology 2007, Web3D 2007, 2007, pp. 61-70. Cited 3 times.

CALLEJAS BEDREGAL, B.R., DIMURO, G.P., ROCHA COSTA, A.C. **Interval fuzzy rule-based hand gesture recognition** (2006) 12th GAMM - IMACS International Symposium on Scientific Computing, Computer Arithmetic and Validated Numerics, SCAN 2006, Conference Post-Proceedings, art. no. 4402402, .

CAMBRA, C., LEAL, A., SILVESTRE, N. **How deaf and hearing adolescents comprehend a televised story.** Deafness and Education International, 12 (1), pp. 34-51. (2010)

CANDIDO, A; MAZIERO, E.; GASPERIN, C.; PARDO T.A.S.; SPECIA, L. ALUISIO S.M. **Supporting the Adaptation of Texts for Poor Literacy Readers: a Text Simplification Editor for Brazilian Portuguese.** Proceedings of the NAACL HLT Workshop on Innovative Use of NLP for Building Educational Applications, pp. 34 - 42, Association for Computational Linguistics. Boulder, Colorado, Junho 2009.

CAÑÓN, L. J. **Accesibilidad a la Información y a la comunicación.** Confederación Estatal de Personas Sordas. Disponível em: <<http://www.cnse.es/comision.php?comision=ACCESIBILIDAD>>. Acesso em: 22 ago. 2005

CAPLIER, A., STILLITTANO, S., ARAN, O., AKARUN, L., BAILLY, G., BEAUTEMPS, D., ABOUTABIT, N., BURGER, T. **Image and video for hearing impaired people.** Eurasip Journal on Image and Video Processing, 2007, art. no. 45641, (2007)

CASELI, H.M.; PEREIRA, T.F., SPECIA, L.; PARDO, T.A.S.; GASPERIN, C.; ALUÍSIO, S.M.; (2009). **Building a Brazilian Portuguese parallel corpus of original and simplified texts.** In Alexander Gelbukh (ed), *Advances in Computational Linguistics, Research in Computer Science*, vol 41, pp. 59-70. 10th Conference on Intelligent Text Processing and Computational Linguistics (CICLing-

2009), March 01–07, Mexico City.

<http://www2.dc.ufscar.br/~helenacaseli/pdf/2009/CICLing2009.pdf>

CASTRO, C. **Uso de Plataformas Tecnológicas para Inclusão Digital - o caso da TV digital e da produção conteúdos**. *Inclusão Social*, Vol. 3, No 1 (2008)

CHABEB, Y., ELGHOUL, O., JEMNI, M.: **WebSign: A Web-Based Interpreter of Sign Language**. In: International Conference on Informatics and its Applications, Oujda (2006)

CHIARO, DELIA. **Issues in audiovisual translation**. Munday, J. *The Routledge companion to translation studies*. pp.141-154. 2009 . Routledge

CHON, J., WHITTLE, S., RISKIN, E.A., LADNER, R.E. **Improving compressed video sign language conversations in the presence of data loss**. *Data Compression Conference Proceedings*, art. no. 5749496, pp. 383-392. (2011)

CIARAMELLO, F.M., KO, J., HEMAMI, S.S.

Quality versus intelligibility: Studying human preferences for American sign language video. In: Western New York Image Processing Workshop. *Proceedings of the WNYIPW 2010*, pp. 70-73, 2010.

CORRADI, JULIANE ADNE MESA E VIDOTTI, SILVANA APARECIDA BORSETTI GREGORIO (2007) **Arquitetura da informação para ambientes informacionais digitais inclusivos: acessibilidade para minorias lingüísticas surdas**. In *Proceedings 7º Encontro Nacional de Ensino e Pesquisa da Informação - CIFORM*, pages 01-14, Salvador (Bahia).

CORRADI, JULIANE ADNE MESA. **Ambientes informacionais digitais e usuários Surdos: questões de acessibilidade**. 214 f. *Dissertação (Mestrado em Ciência da Informação) - Faculdade de Filosofia e Ciências, Universidade Estadual Paulista - UNESP, Marília, 2007.*

CORREIA, ANDERSON TAVARES; MACEDO JR, MÁRCIO RIBEIRO; LIMA, FRANCISCO JOSÉ DE. **O intérprete de Língua**

Brasileira de Sinais no Ensino Fundamental e seu papel na escola comum (Artigo – trabalho de conclusão de curso). Recife: UFPE, Centro de Educação, Coordenação do Curso de Licenciatura Plena em Pedagogia, 2008.

CORREIA, ANDERSON TAVARES; LIMA, ROSÂNGELA A. F.; LIMA, FRANCISCO J. DE. **Datilologia, tradução ou “oralização sinalizada”?**. II Congresso Nacional de Pesquisa em Tradução e Interpretação de Língua de Sinais Brasileira. Florianópolis, 2010.

COOPER, HELEN ; BOWDEN, RICHARD. **Learning Signs from Subtitles: A Weakly Supervised Approach to Sign Language Recognition**. In IEEE CVPR, Miami, FL, USA, June. 2009.

COSTA, A.C.R., DIMURO, G.P.. **A SignWriting-Based Approach to Sign Language Processing. Gesture and Sign Language in Human-Computer Interaction. Lecture Notes in Computer Science, 2002, v. 2298. 2002, p. 202-205.**

DAVIES, S.J.C., AGRAFIOTIS, D., CANAGARAJAH, C.N., BULL, D.R. **Perceptually optimised coding of open signed video based on gaze prediction**. Electronics Letters, v. 43, n. 21, p. 1135-1137, 2007.

DEBEVC, M. , KOSEC, P. , ROTOVNIK, M., HOLZINGER, A.. **Accessible multimodal web pages with sign language translations for deaf and hard of hearing users** (2009) Proceedings - International Workshop on Database and Expert Systems Applications, DEXA, art. no. 5337155, pp. 279-283.

DELGADO, F.U.. **Universal Design for digital television in Spain**. MuTra 2006 – Audiovisual Translation Scenarios: Conference Proceeding, p.1-11. 2006.

DENARDI, R.M.; MENEZES, P.B. ; COSTA, A.C.R.. **AGA-Sign: Animador de Gestos aplicado à Língua de Sinais**. – InfoComp, Journal of Computer Science, v4. n1 , 2005. pp 47 – 53

DENARDI, R.M . **Aga-Sign: Animador de Gestos Aplicado à Língua de Sinais**. Dissertação mestrada em ciência da computação. Universidade Federal do Rio Grande do Sul. 2006

DIZIONARIO. **Dizionario Elettronico di Base Bilingue Lingua Italiana dei Segni-Italiano**. Disponível em <<http://elis.eurac.edu/diz/>>. Acesso em 15 abril 2012.

DREUW, P.; FORSTER, J.; GWETH, Y.; STEIN, D.; NEY, H.; MARTINEZ, G.; LLAHI, J. V.; CRASBORN, O.; ORMEL, E.; DU, W.; HOYOUX, T.; PIATER, J.; MOYA, J. M.; WHEATLEY, M. **SignSpeak – Understanding, Recognition, and Translation of Sign Languages**. 4th Workshop on the Representation and Processing of Sign Languages: Corpora and Sign Language Technologies. 2010.

DREUW, P.; STEIN, D.; DESELAERS, T.; RYBACH, D.; ZAHEDI, M.; BUNGEROTH, J.; NEY, H. **Spoken Language Processing Techniques for Sign Language Recognition and Translation**. Technology and Dissability, v.20, n.2, p. 121–133, junho de 2008.

ELAN - **EUDICO Linguistic Annotator**. Disponível em <<http://www.let.kun.nl/sign-lang/echo/>>. Acesso 02 de fevereiro de 2012.

ELLIOTT, R.; GLAUERT, J.R.W; KENNAWAY, J.R. ; MARSHALL, I. ; SAFAR, E. **Linguistic modelling and language-processing technologies for Avatar-based sign language presentation**. Universal Access in the Information Society, vol. 6, pp. 375-391, 2007.

EFTHIMIOU, E.A; FOTINEA, S.-E.A; VOGLER, C.A; HANKE, T.B; GLAUERT, J.C; BOWDEN, R.D ; BRAFFORT, A.E; COLLET, C.F; MARAGOS, P.G; SEGOUAT, J.H.. **Sign language recognition, generation, and modelling: A research effort with applications in deaf communication**. Lecture Notes in Computer Science, v.5614 LNCS PART 1, p. 21-30, 2009.

EMMOREY, K.; BELLUGI, U. ; KLIMA, E. **Organização neural da língua de sinais. Em Língua de sinais e educação do surdo**. Eds. MOURA, M. C.; LODI, A. C. e PEREIRA, M. C. Sociedade Brasileira de Neuropsicologia. SBNp. São Paulo. 1993.

FAJARDO, I. ; VIGO, M.; SALMERÓN, L.. Technology for supporting web information search and learning in Sign Language. **Interacting with Computers**, v. 21, ago. 2009, pp. 243-256.

FARHADI, A., FORSYTH, D. **Aligning ASL for statistical translation using a discriminative word model** (2006) Proceedings of the IEEE Computer Society Conference on Computer Vision and Pattern Recognition, 2, art. no. 1640930, pp. 1471-1476.

FELICE, M.; MASCIIO, T.; GENNARI, R.. A visual ontology-driven interface for a web sign language dictionary. WISE'07. **Proceedings of the 2007 international conference on Web information systems engineering**. Springer-Verlag Berlin. p.429-440. 2007.

FELS, D.I.; BRANJE, C.; LEE, D.G.; HORNBERG, M.. Emotive Captioning and access to Television. **Proceedings of the Eleventh Americas Conference on Information Systems**, Omaha, NE, USA, p. 2330 – 2337, 2005.

FELS, D.I., GERDZHEV, M., HIBBARD, E., GOODRUM, A., RICHARDS, J., HARDMAN, J. & THOMPSON, N. **Sign language online with signlink studio 2.0**. Lecture Notes in Computer . v. 5616, p. 492-501, 2009.

FERNANDES, M. M.. **A Model for Optimising subtitles for deaf and hard-of-hearing television viewers in South Africa**. Dissertação de mestrado em Artium in English. Potchefstroomse Universiteit. 2003. 104p.

FLOOD, C. M.. **How Deaf and Hard of Hearing Students experience Learning to Write using Signwriting, a Way to Read and Write Signs?**. 300 f. Tese de doutorado em Linguística Educacional, The University of New Mexico, 2002.

GENNARI, R., DI MASCIIO, T. **An ontology for a web dictionary of italian sign language**. Webist 2007 - 3rd International Conference on Web Information Systems and Technologies, Proceedings, WIA, p. 206-213, 2007.

GHOUL, O.E., JEMNI, M. **Multimedia courses generator for deaf children**. International Arab Journal of Information Technology, v.6, n.5, p. 458-463, 2009.

GRIEVE-SMITH, A.B.: **SignSynth: A sign language synthesis application using Web3D and Perl**. In: Wachsmuth, I., Sowa, T. (eds.) 4th International Workshop on Gesture and Sign Language Based Human-Computer Interaction, LNAI 2298. Springer, Heidelberg, pp. 134–145, 2001.

H-ANIM. Humanoid Animation Working Group. **Information technology — Computer graphics and image processing — Humanoid animation (H-Anim)**. Disponível em < <http://www.h-anim.org/>>. Acesso em 26 jul 2011.

HANKE, THOMAS. **HamNoSys – Hamburg Notation System for Sign Languages**. Apresentação disponível em <http://www.sign-lang.uni-hamburg.de/dgs-korpus/tl_files/inhalt_pdf/HamNoSys_06en.pdf>. Acesso em 01 de abril de 2011.

HAVASI, L.; SZABÓ, H.M.. **A Motion Capture System for Sign Language Synthesis: Overview and Related Issues**. pp 636 – 641. In Advances in Visual Computing. Springer Berlin / Heidelberg 2005

HELOIR, A.; GIBET, S.. **A qualitative and Quantitative Characterization of Style in Sign Language Gestures**. Em Gesture in Human-Computer Interaction and Simulation, 7th International Gesture Workshop, GW 2007, 2007

OKURA, T.; HIROSE, Y.. **Development of a web type DVD viewer synchronized with multilingual captions for existing DVDs**. Lecture Notes in Computer Science, n.5105, pp. 640-646. 2008.

HOLDEN EUN-JUNG, OWENS ROBYN: **Visual sign language recognition; in Multi-Image Search and Analysis**. Klette, R., Huang T., and Gimel'farb, G. (Eds) Lecture Notes in Computer Science, n. 2032, Springer 2001; pp. 270-287

HONG, P.; TURK, M.; HUANG, T.S. **Gesture modeling and recognition using finite state machines**. Beckman Inst., Illinois Univ., Urbana, IL; Automatic Face and Gesture Recognition, 2000. Proceedings. Fourth IEEE International Conference on, Publication. pp. 410-415, Grenoble, France. 2000

HOOPER, S., MILLER, C., ROSE, S., VELETSIANOS, G. **The effects of digital video quality on learner comprehension in an American sign language assessment environment.** Sign Language Studies, n.8, v. 1, p. 42-55, 2007.

HUENERFAUTH, M., 2008. **Generating American Sign Language animation: overcoming misconceptions and technical challenges.** Universal Access in the Information Society 6 (4), 419–434.

HUENERFAUTH, M. **American Sign Language Spatial Representations for an Accessible User-Interface.** In 3rd International Conference on Universal Access in Human-Computer Interaction. Las Vegas, NV, USA. 2005.

IBGE 2010 - **Guia do Censo 2010 para jornalistas.** Disponível em <http://www.ibge.gov.br/home/presidencia/noticias/guia_do_censo_2010_questionarios.php>. Acesso em 10 jan 2011.

IEEE. Taxonomy, Version 1.01. The Institute Of Electrical And Electronics Engineers. 2009.

IGUMA 2004 - "**Surdos e a comunicação audiovisual: desafiando barreiras**". Dissertação

ITU-T. TELECOMMUNICATION STANDARDIZATION SECTOR OF INTERNATIONAL TELECOMMUNICATION UNION. Application profile – Sign language and lip-reading real-time conversation using low bit-rate video communication. SERIES H: AUDIOVISUAL AND MULTIMEDIA SYSTEMS. 1999.

ITU, H.264 Advanced video coding for generic audiovisual devices. International Telecommunications Union. Disponível em <<http://www.itu.int/rec/T-REC-H.264/e>>. Acesso e 15 mar 2012.

JAKOBSON, R. **Aspectos linguísticos da tradução.** Trad. Izidoro Blikstein. In: JAKOBSON, R. Linguística e Comunicação. São Paulo: Cultrix, 1995, p. 63-86.

JANUÁRIO, G. C.; LEITE, L. A. F.; KOGA, M. L.. **POLI-LIBRAS UM TRADUTOR DE PORTUGUÊS PARA LIBRAS.** Monografia

(Engenharia da Computação). Escola Politécnica da Universidade de São Paulo. 2010.

JATOBÁ, R. **Previsão do tempo do Jornal Nacional**. Julho de 2008. Disponível em <<http://www.youtube.com/watch?v=ImcABBOuETc>>. Acesso em 10 jan 2010.

JEMNI, M. ; ELGHOUL, O. **A System to Make Signs Using Collaborative Approach**. Proceedings of the 11th international conference on Computers Helping People with Special Needs, Inz, Austria: Springer-Verlag, 2008, pp. 670-677.

JEMNI, M., ELGHOUL, O. **Towards Web-Based automatic interpretation of written text to Sign Language**. In: First International conference on ICT & Accessibility, Tunisia (2007) a

JEMNI, M., ELGHOUL, O., MAKHLOUF, S.: **A Web-Based Tool to Create Online Courses for Deaf Pupils**. In: IMCL Conference, Amman (2007)

JISC - Joint Information Systems Committee, Digital Media - Institute for Learning and Research Technology (ILRT) at the University of Bristol. Disponível em <<http://www.jiscdigitalmedia.ac.uk/movingimages/advice/adding-sign-language-to-a-video/>>. Acesso em out de 2010 .

JOHNSTON, T., NAPIER, J. **Medical signbank: Bringing deaf people and linguists together in the process of language development** (2010) Sign Language Studies, 10 (2), pp. 258-275.

JUNIOR, E.B.S.. **Convergência Digital para Apoio ao Ensino de Libras, com ênfase na Web e no Sistema Brasileiro de TV Digital**. Universidade Estadual Paulista “Júlio de Mesquita Filho”. Dissertação de mestrado. Programa de Pós-Graduação em Ciência da Computação. São José do Rio Preto, SP. 2011. 106p.

KADOUS, WALEED. **GRASP - Recognition of Australian Sign Language using Instrumented Gloves**; University of New South Wales , Australia, Oct 1995.

KANEKO, H., HAMAGUCHI, N., DOKE, M., INOUE, S.
Sign language animation using TVML. In: Proceedings - VRCAI 2010, ACM SIGGRAPH Conference on Virtual-Reality Continuum and Its Application to Industry, p. 289-292, 2010.

KENNAWAY, R. (2004). **Experience with and requirements for a gesture description language for synthetic animation.** In: Lecture Notes in Computer Science Vol. 2915, pp. 300-311

KENNAWAY, J.R., GLAUERT, J.R.W., ZWITSERLOOD, I.
Providing signed content on the Internet by synthesized animation (2007) ACM Transactions on Computer-Human Interaction, 14 (3), art. no. 15,

KIM, Y.; YANG, J.; PARK, M.; AHN, S.; AHN, C.. **Interactive broadcast terminal system using MPEG-2 and MPEG-4 Circuits and Systems**, 2000. Proceedings. ISCAS 2000 Geneva. The 2000 IEEE International Symposium on Circuits and Systems, Publication Date: 2000, v.3, p. 682-685, Geneva, Switzerland, 2000.

KOUNTCHEV, R., TODOROV, VL., KOUNTCHEVA, R. **Efficient sign language video representation.** Proceedings of IWSSIP 2008 - 15th International Conference on Systems, Signals and Image Processing, art. no. 4604396, pp. 177-180, 2008.

KRAPEŽ, S., SOLINA, F. **Synthesis of the sign language of the deaf from the sign video clips.** (1999) Elektrotehniski Vestnik/Electrotechnical Review, 66 (4-5), pp. 260-265.

KURODA, T; TABATA, Y; GOTO, A.; IKUTA, H.; MURAKAMI, M;
Consumer price data-glove for sign language recognition. Proc. 5th Intl Conf. Disability, Virtual Reality & Assoc. Tech., Oxford, UK, 2004.

KURZ, I.; B. MIKULASEK. **Television as a Source of Information for the Deaf and Hearing Impaired. Captions and Sign Language on Austrian TV.** Meta, v. 49, n. 1, p.81-88, 2004.

LAMBOURNE, A.; HEWITT, J.; LYON, C.; WARREN, S. **Speech-Based Real-Time Subtitling Services.** International Journal of Speech Technology, v.7, n. 4, p. 269-279, 2004.

LEELARASMEE, E., MAHASITH, R., UTHAICHANA, P. **A novel video sign image expander**. In: Proceedings of the International Symposium on Consumer Electronics, ISCE, p. 391-394, 2005.

LEWIS, M. S. J.; JACKSON, D. W.. **Television Literacy: Comprehension of Program Content Using Closed Captions for the Deaf**. *J. Deaf Stud. Deaf Educ.* v.6, n.1, p.43-53, 2001.

LICHTIG, I.; COUTO M.; LEME V.N.. **Perfil pragmático de crianças surdas em diferentes fases lingüísticas**. *Revista da Sociedade Brasileira de Fonoaudiologia*, v. 13, n.3, p.251-7, 2008.

LIRA, G.A. **O Impacto da Tecnologia na Educação e Inclusão Social da Pessoa Portadora de Deficiência Auditiva: Tlibras Tradutor Digital Português x Língua Brasileira de Sinais – Libras**. Disponível em <<http://www.senac.br/BTS/293/boltec293d.htm>>. Acesso em 10 março de 2009.

LU, CHING-TING. **Analysis of English subtitles produced for the Taiwanese movie Cape No.7**. Dissertação de mestrado. Master of Arts in Applied Language Studies. AUT University. 2010. 144p.

MARTIN, C.A.; GARCIA, L.; MENENDEZ, J.M.; CISNEROS, G.; , **Access services based on MHP interactive applications**," *Consumer Electronics (ICCE), 2010 Digest of Technical Papers International Conference on* , vol., no., pp.459-460, 9-13 Jan. 2010

MASUTTI, MARA LÚCIA. **TRADUÇÃO CULTURAL: DESCONSTRUÇÕES LOGOFONOCÊNTRICAS EM ZONAS DE CONTATO ENTRE SURDOS E OUVINTE**. Tese apresentada como requisito parcial à obtenção do grau de Doutor no Curso de Pós-Graduação em Literatura do Centro de Comunicação e Expressão da Universidade Federal de Santa Catarina. 2007. 157p

MALDONADO, JAVIER MELENCHÓN. **Síntesis Audiovisual Realista Personalizable**. Tese de doutorado. Universitat Ramon Liull. EALS – Comunicacions I Teoria del Senyal. 199p.

MAURER, HERMANN; STUBENRAUCH, , ROBERT; CAMHY, DANIELA G. **Foundations of MIRACLE: Multimedia Information**

- Repository.** A Computer-supported Language Effort Journal of Universal Computer Science, v. 9, n. 4, 2003.
- MCCARTY, A.L.. Notation Systems for Reading and Writing Sign Language . The Analysis of Verbal Behavior, n. 20, p. 129-134, 2004.
- MELLO, A. G. **Surdos Oralizados e Não Oralizados: Uma Visão Crítica.** In: SEGUNDO CONGRESSO VIRTUAL “Integración sin Barreras em el Siglo XXI”. Anais eletrônicos... Red de Integración Especial, 2001. Disponível em <<http://www.redespecialweb.org>>. Acesso em: fev. 2002.
- MILLER, C. Access symbols for use with vídeo content and information and communications technology devices. In: DÍAZ, J.C.; CLAVERO, P.O.; REMAEL, A.. Media for all: subtitling for the deaf, audio description, and sign language. Amsterdam: Rodopi. 2007. p. 53-69.
- MITOBE, K., et al. 2006. **Development of a Motion Capture System for a Hand Using a Magnetic Three Dimensional Position Sensor,** ACM SIGGRAPH2006, Research Posters, 115
- MÖBUS, L. **Making web content accessible for the deaf via sign language.** Library Hi Tech, v. 28 ,n.4, p. 569-576. 2010.
- MOEMEDI, K.; CONNAN, J.. **Rendering an Animated Avatar from SignWriting Notation.** Proceedings of SATNAC conference. 2010.
- MONTEZ, C.; BECKER, V.. **TV Digital Interativa: conceitos, desafios e perspectivas para o Brasil.** Florianópolis: Ed. da UFSC, 2005. 2ª edição.
- MOTTA, R. F.. **Sociedade Inclusiva: O Acesso do Surdo à TV** Regiocom. Chapeco, 2005.
- MOODLE. **Modular Object Oriented Learning Environment.** Disponível em <<http://www.moodle.org>>. Acesso em 14 abril 2012.
- MUIR, L. J.. **Content-prioritised video coding for British Sign Language communication.** Tese de doutorado em . The Robert Gordon University, 2007. 269 p.

MUIR, L. J.; RICHARDSON, I.E. G.. **Video telephony for the deaf: Analysis and development of an optimized video compression product**. In: Proc. of Tenth ACM International Conference on Multimedia, p.650–652, 2000.

_____. Perception of sign language and its application to visual communications for deaf people. *Journal of Deaf Studies and Deaf Education*. Oxford University Press, v.10, n.4, pp. 390-401, 2005.

NASCIMENTO, M.V.B. **Interpretação da Libras no gênero jornalístico televisivo: elementos extralinguísticos na produção dos sentidos**. III Congresso Nacional de Pesquisas em Tradução e Interpretação de Libras e Língua Portuguesa. 2010. Disponível em <http://www.congressotils.cce.ufsc.br/2010/pdf/marcus_vinicius_batista_nascimento.pdf>. Acesso em 12 jan 2011.

NASCIMENTO, M. V. B. **Interpretação da língua brasileira de sinais a partir do gênero jornalístico televisivo: elementos verbos-visuais na produção de sentidos**. Dissertação de Mestrado em Linguística Aplicada e Estudos da Linguagem. PUC –SP. São Paulo, 2011. 147p.

NAKAZONO, K., NAGASHIMA, Y., ICHIKAWA, A. **Digital encoding applied to sign language video**. *IEICE Transactions on Information and Systems*, v. E89-D, n.6, p. 1893-1900, 2006.

NEVES, J.. **Audiovisual Translation: Subtitling for the Deaf and Hard-of-Hearing**. Tese de doutorado. School of Arts, Roehampton University. University of Surrey, Londres, 2005, 358 p.

NEVES, J.. **Vozes que se Vêm – Guia de Legendagem para Surdos**. Leiria: Instituto Politécnico de Leiria & Universidade de Aveiro. 100 p. 2007.

OASIS, S. Reference Model. OASIS SOA Technical Committee. Disponível em <<http://www.oasis-open.org/committees/soa-rm/faq.php>>. Acesso em 14 de Nov de 2011.

OBREGON, R. F. A.. **O Padrão Arquétipo da Alteridade e o Compartilhamento de Conhecimento em Ambiente Virtual de**

Aprendizagem Inclusivo. Tese de doutorado. Programa de Engenharia e Gestão do Conhecimento. Florianópolis. UFSC. 2011. 208p.

OLIVEIRA, J.S. **Glossário Letras-Libras como Ferramenta para Formação/Consulta de Tradutores.** III Congresso Nacional de Pesquisas em Tradução e Interpretação de Libras e Língua Portuguesa. Anais do III Congresso Nacional de Pesquisas em Tradução e Interpretação de Libras e Língua Portuguesa. Florianópolis, 2010.

OI, M. F.; TANG, L.; WAI, L.. **V2S: Voice to Sign Language Translation System for Malaysian Deaf People.** In Visual Informatics: Bridging Research and Practice. Lecture Notes in Computer Science, 2009. Springer Berlin / Heidelberg. p. 868-876. v. 5857.

ONG, S.C.W., RANGANATH, S. **Automatic sign language analysis: a survey and the future beyond lexical meaning .** IEEE Transactions on Pattern Analysis and Machine Intelligence, 27 (6), pp. 873-891. 2005.

OPENGINA. Ginga Code Development Network. Disponível em <<http://gingacdn.lavid.ufpb.br/projects/openginga/>>. Acesso em 28 abril de 2012.

OTHMAN, A., GHOUL, O.E., JEMNI, M. **SportSign: A Service to Make Sports News Accessible to Deaf Persons in Sign Languages.** In Proceedings of ICCHP. Lecture Notes in Computer Science, v. 6180, p. 169-176, 2010.

PAPADOGIORGAKI, M.; GRAMMALIDIS N.; SARRIS, N.; STRINTZIS, M.G.. **Synthesis of virtual reality animation from sign language notation using MPEG-4 body animation parameters.** Proc. 5th Intl Conf. Disability, Virtual Reality & Assoc. Tech., Oxford, UK, 2004.

PAPADOGIORGAKI, M., GRAMMALIDIS, N., MAKRIS, L. **Sign synthesis from Sign Writing notation using MPEG-4, H-Anim, and inverse kinematics techniques.** Journal of Endocrine Genetics, v. 4, n. 3, p. 191-203, 2005.

PARKHURST, S.; PARKHURST, D.. A Cross-Linguistic Guide to SignWriting. A phonetic approach. 152 p. 2010.

PETTIT, Zoë. The Audio-Visual Text: Subtitling and Dubbing Different Genres. Special Issue on Audiovisual Translation. *Meta*. v. 49, n.1, pp. 25-38. 2004.

PHAN, DA PHUC ; NGUYEN, THI NHAT THANH ; BUI, THE DUY. **A 3D Conversational Agent for Presenting Digital Information for Deaf People**. Agent Computing and Multi-Agent Systems. Lecture Notes in Computer Science, 2009, Volume 5044/2009, 319-328,

PICCOLO, L.; MELO, A.; BARANAUSKAS, M.. **Accessibility and interactive tv: Design recommendations for the brazilian scenario**. In Human-Computer Interaction – INTERACT 2007. 11th IFIP TC 13 International Conference, Rio de Janeiro, Brazil, September 10-14, 2007, Proceedings, Part I. Lecture Notes in Computer Science. pages 361–374. 2007

PISTORI, H. ; MARTINS, P. S. ; PEREIRA, M. C. ; NETO, J. J.. **Sigus - plataforma de apoio ao desenvolvimento de sistemas para inclusão digital de pessoas com necessidades especiais**. IV Congresso Iberdiscap: Tecnologias de Apoio a Portadores de Deficiência, Vitória, February 2006.

PISTORI, H. **Computer vision and digital inclusion of persons with special needs: Overview and state of art**. Proceedings of the International Symposium CompIMAGE 2006 - Computational Modelling of Objects Represented in Images: Fundamentals, Methods and Applications, pp. 33-37. 2007.

POLI-LIBRAS. **Um Tradutor Português-Libras**. Disponível em <<http://www.polilibras.com.br/>>. Acesso em: 11 jul 2011.

POLLARD JR., R.Q., DEAN, R.K., O'HEARN, A., HAYNES, S.L. **Adapting Health Education Material for Deaf Audiences**. *Rehabilitation Psychology*, 54 (2), pp. 232-238. 2009.

POURTOIS, JEAN-PIERRE; DESMET, HUGUETTE. **Épistémologie et instrumentation en sciences humaines**. Liège: Mandaga, 1997.

PREDA, M; PRETEUX, F. **Insights into low-level avatar animation and MPEG-4 standardization.** Signal Processing: Image Communication, Volume 17, Issue 9, October 2002, Pages 717-741.

PUC. The Programming Language Lua. Disponível em <<http://www.lua.org/>> . Acesso em 15 jan 2012.

QUADROS, R. M. DE; KARNOPP, L. B.. **Língua de sinais brasileira: estudos linguísticos.** Porto Alegre: Artmed, 2004.

QUADROS, R.M.. O tradutor e intérprete de língua brasileira de sinais e língua portuguesa. Brasília. 2004. Secretaria de Educação Especial. Ministério da Educação. 94p.

RANDER, A.; LOOMS, P. O. 2010. **The accessibility of television news with live subtitling on digital television.** In Proceedings of the 8th international interactive Conference on interactive Tv&Video (Tampere, Finland, June 09 - 11, 2010). EuroITV '10. ACM, New York, NY, 155-160. DOI= <http://doi.acm.org/10.1145/1809777.1809809>

RED5. Red5 The Open Source Media Server. Disponível em <<http://www.red5.org/>>. Acesso em 28 mai de 2012.

REICHERT. ANDRÉ RIBEIRO. **Mídia televisiva sem som.** Dissertação de mestrado em educação. Programa de Pós-Graduação em Educação. Faculdade de Educação. Universidade Federal do Rio Grande do Sul. Disponível em <http://www.lume.ufrgs.br/handle/10183/10016>. 2006. 100 p.

RIEGER , T; BRAUN, N. **Narrative Use of Sign Language by a Virtual Character for the Hearing Impaired,** Eurographics vol. 22, n. 3, pp. 651- 660, 2003.

ROCHA, T. C. da C.. DICIONÁRIO TEMÁTICO DA LÍNGUA BRASILEIRA DE SINAIS: A Criação de Sinais Específicos da Filosofia. XXIII Prêmio Jovem Cientista Educação para Reduzir as Desigualdades Sociais. 18p. Pontifícia Universidade Católica de Minas Gerais. Belo Horizonte. 2008.

ROBINSON, L.. CVP & GM, Home Devices, Motorola Mobility. 23/09/2011. Disponível em <<http://www.youtube.com/watch?v=n9wc1y-wU68>>. Acesso em 15 nov 2011.

RODRIGUES, N. **Organização neural da linguagem. Em Língua de sinais e educação do surdo.** Eds. Moura, M. C.; LODI, A. C. e PEREIRA, M. C. Sociedade Brasileira de Neuropsicologia. SBNp. São Paulo. 1993.

RYBENA, Software. Disponível em: <<http://www.rybena.com.br/>>. Acesso em mai de 2010.

SAN-SEGUNDO, R., BARRA, R., CÓRDOBA, R., D'HARO, L.F., FERNANDEZ, F., FERREIROS, J., LUCAS, J.M., PARDO, J.M.. **Speech to sign language translation system for Spanish.** Speech Communication. n.50 , v.11-12, p. 1009-1020. 2008.

SAN-SEGUNDO, R. ; MONTERO, J.M.; MACIAS-GUARASA, J.; CORDOBA, R.; FERREIROS, J.; PARDO, J.M.. **Proposing a speech to gesture translation architecture for Spanish deaf people.** Journal of Visual Languages & Computing, v. 19, n. 5, outubro 2008, p. 523-538.

SANTOS, E.B. JR.; OLIVEIRA H.C.; OLIVEIRA, C.S.. **Acessibilidade na TV Digital Aberta no Brasil para apoio a surdos.** Proceedings of I2TS 2010 – 9th International Information and Telecommunication Thecnologis Synposium. Rio de Janeiro, 2010.

SANTOS, J. F.M.. **DANDO VOZ AO SURDO: A inclusão escolar na visão do aluno surdo.** Graduação em Fonoaudiologia. . Disponível em <<http://www.nupad.ufmg.br/fon/monografias/20082/Janice.pdf>> . Acesso em 15 dez 2008.

SCHELLING, THOMAS C. , **Micromotives and Macrobehavior.** New York: Norton. 1978.

SEGALA, RIMAR RAMALHO. **Tadução Intermodal e Intersemiótica/Interlingual: Português brasileiro para Língua Brasileira de Sinais.** Dissertação de mestrado. Pós-Graduação em Estudos da Tradução. Universidade Fderal de Santa Catarina. Março de 2010. 74p.

SELVATICI, C.. **Closed caption: conquistas e questões.** Dissertação de Mestrado. Programa de Pós Graduação em Letras da PUC-Rio. Jan 2010. 870p.

SHOHIEB, S.M.A; HASSAN, A.E.C; ELSOUD, M.A.B; KANDIL, M.S.C.. **Accessibility system for deaf Arab students.** Human Capital Development for Progress: 2009 ITI 7th International Conference on Information and Communications Technology, ICICT 2009.

SILVA, A.B.P.; PEREIRA, M.C.C.; ZANOLLI, M.L.. **Mães ouvintes com filhos surdos: concepção de surdez e escolha da modalidade de linguagem.** Psic.: Teor. e Pesq. 2007, v.23, n.3, p. 279-286. Brasília, Jul-Set 2007.

SILVA, E. L.; MENEZES, E.M.. **Metodologia da Pesquisa e Elaboração de Dissertação.** Terceira edição. Florianópolis. Laboratório de Educação a Distância. 2001. 118p.

SILVA, F.I.. **Analisando o processo de leitura de uma possível escrita da língua brasileira de sinais: Signwriting.** 114 p. Dissertação (Mestrado em educação). UFSC. Florianópolis, SC, 2009.

SILVA, V. **A luta dos Surdos pelo direito à Educação e ao Trabalho: Relato de uma Vivência Político-Pedagógica na Escola Técnica Federal de Santa Catarina.** 142 p. Dissertação (Mestrado em Educação). UFSC. Florianópolis, SC, 2001.

SIGNBANK.**Binary Sign Writing HTML Reference.** Disponível em < <http://www.signbank.org/bsw/> >. Acesso em 15 mai 2012.

SIGNSYNTH. Um protótipo articulatório de língua de sinais. Disponível em <<http://www.panix.com/~grvsmth/signsynth/>>. Acesso em 14 jul 2011.

SIGNTEXT, Disponível em < <http://www.signbank.org/signtext/> >.Acesso em 10 jan 2012.

SIGN WRITER STUDIO. Disponível em <<http://signwriterstudio.com/download.htm>>. Acesso em 10 jan 2012.

SOUZA, V.C.; PINTO, S.C.. **Sign WebMessage: uma ferramenta para comunicação via web através da Língua Brasileira de Sinais – Libras.** In: XIV Simpósio Brasileiro de Informática na Educação – SBIE. RJ, 2003.

SOUZA, V.C.. **SWSservice: uma biblioteca para a escrita da Língua Brasileira de Sinais baseada em Web Services.** 129 f. Dissertação (Mestrado em Computação Aplicada). Programa Interdisciplinar de pós-graduação em Computação Aplicada. Universidade do Vale do Rio dos Sinos. 2005. (b)

SOARES, L.F.G.; MORENO, M.F.; COSTA, R.M.R.C; MORENO, M.F. **Towards the convergence of digital TV systems.** Journal of Internet Services and Applications, v.1 n.1. pp 69–79, 2010. Springer

SOUZA, J.C.T.. **Ouidos silenciados, mãos que falam: os surdos e a teleinformação.** 141 f. Dissertação (Mestrado em Educação). Programa de pós-graduação em Educação, UFBA. 2005.

SOUZA, K. P.; PISTORI, H.. **Implementação de um Extrator de Características baseado em Momentos da Imagem.** XVIII Brazilian Symposium on Computer Graphics and Image Processing - SIBGRAPI, III Workshop de Trabalhos de Iniciação Científica em Computação Gráfica e Processamento de Imagens - WICCGPI, Natal, Outubro 9-12, 2005.

STANOEVSKA-SLABEVA, K.. **The Concept of Knowledge Media: The Past and Future.** In Knowledge Media in Healthcare: Opportunities & Challenges. Idea Group Publishing. Grütter Rolf. USA. 2002. 283p.

STOKOE, WILLIAM C. 1960. **Sign Language Structure.** Studies in Linguistics. Occasional Paper 8. Buffalo: University of Buffalo Press.

STOKOE, WILLIAM C. **Language in Sign: why sign came before speech.** 2001 Gallaudet University Press, 256p.

STOKOE, WILLIAM C. ET al. 1965. **A Dictionary of American Sign Language on Linguistic Principles.** Washington, D.C.: Gallaudet College Press.

STUMPF, M. R. **Língua de sinais: escrita dos surdos na internet**. In RIBIE - V Congresso Iberoamericano de Informática Educativa, 2000, Viña del Mar - Chile, 2000.

SUTTON, V.; GLEAVES, R.. **SignWriter - The world's first sign language processor**. La Jolla, CA: Ed. Center for Sutton Movement Writing, 1995.

SUTTON, V.. **Lições sobre o SignWriting Um Sistema de Escrita para Língua de Sinais**. Tradução Parcial e Adaptação do Inglês/ASL para Português LIBRAS do livro “Lessons in SignWriting “, de Valerie Sutton, publicado originalmente pelo DAC – Deaf Action Committe for SignWriting. Tradução: Marianne Rossi Stumpf . 2008.

SUTTON, V.. **Sign writing**, 1990. Disponível em: <http://www.signwriting.org>. Acessado em set, 2000.

SUTTON, V.. **Sutton's SignSpelling Guidelines 2008** , Center for Sutton Movement Writing, Manual disponível em www.signwriting.org/archive/.../sw0534-SignSpellingGuidelines-2008.pdf) SignWriting Press) (2008)

SUTTON, V.; FROST, A.. **SignWriting Reference Manual**. Part 2: SignWriting Hand Symbols. The SignWriting Press. 2011. 310p.

SW-EDIT. **Editor de textos e para línguas de sinais**. Disponível em < http://rocha.c3.furg.br/index.php?Itemid=513&option=download_categoria&task=detalhe&id_site_componente=707&id=31> . Acesso em 31 jan de 2012.

TANIA, D.M.; MAURO, F.; ROSELLA, G.. **Visually querying the e-LIS ontology: A proposal**. In: Proceedings of Conference on Human Factors in Computing Systems, p. 3597-3602, 2008.

TAVARES, J.E.R.; BARBOSA, J.L.V.; LEITHARDT, V.R.Q.. **SENSORLIBRAS: Tradução Automática Libras-Portugues através da Computação Ubiqua**. II Congresso Nacional de pesquisa em Tradução e Interpretação de Língua de Sinais Brasileira. Nov 2010.

TAVARES, O. L.; CORADINE, L. C. ; BREDA, W.L.; **Falibras-MT – Autoria de tradutores automáticos de textos do português para**

LIBRAS, na forma gestual animada: Uma abordagem com memória de tradução. Anais do XXV congresso da Sociedade brasileira da computação. 2005.

TEIXIDOR, E. B.. **PARÀMETRES PER A UNA TAXONOMIA DE LA SUBTITULACIÓ.** 449f. Tese (Doutorado em tradução). Departamento de Traducció i Filologia. Universidade Pompeu Fabra. 2008.

TORREZ, E.F.; MAZZONI, A. A.; MELLO, A. G.. **Nem toda pessoa cega lê em Braille nem toda pessoa surda se comunica em língua de sinais.** Educação e Pesquisa. 2007, v. 33, n. 2, p. 369-386.

TORREZ, E.F.; MAZZONI, A. A.; ALVES, J.,B.M.. **The accessibility to the information in the digital space.** Ci. Inf., Brasília, v. 31, n. 3, setembro 2002 .

TRAN, J.J.; JOHNSON, T.W.; KIM, J.; RODRIGUEZ, R.; YIN, S.; RISKIN, E.A.; LADNER, R.E.; WOBROCK, J.O.. **A web-based user survey for evaluating power saving strategies for deaf users of MobileASL.** In: Proceedings of the 12th International ACM SIGACCESS Conference on Computers and Accessibility, p. 115-122. 2010.

TTML. **Timed Text Markup Language.** Versão 1.0. W3C Recommendation. Disponível em <<http://www.w3.org/TR/ttaf1-dfxp/>>. Acesso em nov 2010.

TVEIT, J. E.. **Translating for Television, A Handbook in Screen Translation.** Bergen: JK Publishing. 139p., 2004.

UDO, J.P., FELS, D.I. The rogue poster-children of universal design: Closed captioning and audio description. *Journal of Engineering Design*, v. 21, n. 2-3, pp. 207-221, 2010.

UGARTE, A., GARCIA, I., ORTIZ, A., OYARZUN, D. User Interfaces Based on 3D Avatars for Interactive Television. *Lecture Notes in Computer Science.* Springer Berlin / Heidelberg. v. 4471, pp. 107-115, 2007.

URUSOFT. **Subtitle Workshop**. Disponível em <<http://www.urusoft.net/products.php>>. Acesso em abr 2012.

VEIGA, E. G; TAVARES, T. A., Um Modelo de Processo para o Desenvolvimento de Programas para TV Digital e Interativa baseado em Metodologias Ágeis. In: **Primeiro Workshop em Desenvolvimento Rápido de Aplicações**, 2007, Porto de Galinhas, Pernambuco.

VERLINDEN, M.; ZWITSERLOOD, I.; FROWEIN, H. **Multimedia with Animated Sign Language for Deaf Learners**. In: Proceedings of ED-Media. p. 4759-4764, 2005.

VHML. **Virtual Human Markup Language: Working Draft v0.3**. Disponível online em <http://www.vhml.org/documents/VHML/2001/WD-VHML-20011021/> Acesso em: 15 nov de 2008

VISICAST. **Virtual Human Signing at UEA**. University of East Anglia. Disponível em <<http://www.visicast.cmp.uea.ac.uk>>. Acesso em 14 jun 2011.

VSIGNS. Informatics and Telematics Institute, Grécia, 2004. Disponível em <<http://vsigns.iti.gr>>. Acessado em abr 2009.

VY, Q.V.; FELS, D.I. Using avatars for improving speaker identification in captioning. **Lecture Notes in Computer Science**, v. 5727, n. 2, p. 916-919, 2009.

VY, Q.V., FELS, D.I. **Using placement and name for speaker identification in captioning**. **Lecture Notes in Computer Science** (including subseries Lecture Notes in Artificial Intelligence and Lecture Notes in Bioinformatics) 6179 LNCS (PART 1), p. 247-254. 2010

WALD, M.. **Captioning for deaf and hard of hearing people by editing automatic speech recognition in real time**. **Lecture Notes in Computer Science**, v. 4061, p. 683-690, 2006.

WANDERLEY, D.. Fonologia em Escrita de Sinais. Trabalho de conclusão da disciplina de Aquisição de Língua de Sinais. Prof. Ronice M. de Quadros. UFSC. 2011.

WANG, Q.; CHEN, X.; ZHANG, L.; CHUNLI, W.; GAO, W..
Viewpoint invariant sign language recognition. Computer Vision and Image Understanding. v. 108 , n. 1-2, p. 87-97. 2007.

Wiki-Libras. Disponível em
<<http://www.polidada.poli.usp.br/wikilibras/>>. Acesso em 12 jan 2012.

WILLIAMS, H.; THORNE, D.. **The value of teletext subtitling as a medium for language learning,** System, Volume 28, Issue 2, June 2000, pp. 217-228, 2000.

WÖHRMANN, S.. **Learn to read SignWriting (GebaerdenSchrift).** Das Zeichen Journal, 2003. Disponível em
<<http://signwriting.org/germany/DasZeichen/>>. Acesso em 15 jan 2011.

XU, X., WANG, X., YAO, L., ZHANG, D., ZHAO, H. **Avatars based chinese sign language synthesis system.** In: Conference Proceedings of IEEE International Conference on Systems, Man and Cybernetics. pp. 444-449, 2008.

YANAGIMOTO, M.,. Sign Language Animation Site “Hello! Astroboy”. **ACM SIGGRAPH 2004.** ACM Press. 2004.

ZHANG, YONGJUN. **A Java 3D Framework for Digital Television Set-top Box.** 57 p. Dissertação (Mestrado em Engenharia e Ciência da Computação). Helsinki University of Technology, 2003.

ZAPPE, C. T.. **Escrita da Língua de Sinais em comunidades do Orkut: marcador cultural na educação de surdos.** 68 f. Dissertação (Mestrado em Educação). Pós Graduação em Educação. Universidade Federal de Santa Maria, 2010.

ZWITERSLOOD, I., VERLINDEN, M., ROS, J., SCHOOT, S. Synthetic signing for the deaf: Esign. In: **Conference and Workshop on Assistive Technologies for Vision and Hearing Impairment,** Granada, Spain. Julho de 2004.

ANEXO A – REVISÃO SISTEMÁTICA DA BIBLIOGRAFIA

O universo de estudo foi elaborado sobre os documentos disponíveis nas bases de periódicos, dissertações, teses e publicações disseminadas nas Bibliotecas Nacional de Teses e Dissertações, *Networked Library of Thesis and Dissertations*, além da bases de periódicos *Scopus* e *ScienceDirect*. Além dessas bases foi utilizado o recurso de busca *Google Scholar* para localizar artigos referenciados nos documentos analisados.

Descritores de palavras-chave foram selecionados a partir da taxonomia do IEEE (2009). Dessa taxonomia partiram os termos [Deafness], que caracteriza os publico a ser beneficiado pela pesquisa e [TV], [web] e [internet], que são os meios de publicação do audiovisual.

A partir da definição dos termos da pesquisa, a seguinte expressão foi usada nas buscas nas bases de documentos:

(deaf OR deafness OR sign language) AND (television OR TV OR DTV OR IDTV OR captions OR legenda OR subtitles OR digital OR broadcast OR web OR audiovisual OR vídeo OR movie)

As artigos científicos, dissertações e teses foram analisados e selecionados pela leitura de títulos, resumos e corpo de textos dos documentos encontrados. Especificamente foram selecionadas as publicações que relacionem a surdez ao audiovisual, focando em aspectos como o desenvolvimento de legendas e preferências de surdos em relação aos recursos de acessibilidade.

Foram obtidos 975 resultados na base Scopus, dos quais todos foram analisados e 46 referenciados. Já na base Science Direct, foram encontrados 39.878 documentos dos quais os 500 mais relevantes foram analisados e apenas 4 foram referenciados (em 18/06/2011).

Após a busca e respeitando os critérios de inclusão, foram encontrados 50 artigos, 5 teses e 11 dissertações. Sendo distribuídos de acordo com os quadros 1, 2 e 3 a seguir.

Quadro 16: quantidade de artigos resultantes da revisão sistemática

	Artigos	
Base	Analisados	Referenciados

Scopus	975	46
Science Direct	500	4

Quadro 17: quantidade de teses e dissertações resultantes da revisão sistemática

Base	documentos analisados	teses referenciadas	dissertações referenciadas
NLTD	500	4	4
BNTD	137	1	7

Seguem os resultados utilizados de ScienceDirect e suas respectivas descrições:

- Fajardo, Vigo, Salmerón (2009) - taxonomia de tecnologias de suporte;
- Adamo-Villani, Beni (2005) - legendagem em língua de sinais;
- Preda, Preteux (2002) - animação de avatares com FAP e BAPs sobre mpeg TS;
- San-Segundo et al. (2008) - reconhecimento de voz e apresentação por agente de representação de gestos.

Os resultados de Scopus e descrições:

- Krapež, S., Solina (1999) - síntese de vídeo em LS a partir da junção de vídeo clipes;
- Allen, J.M., Asselin, P.K., Foulds, R. (2003) - reconhecimento de datilologia dos sinais;
- Bain, Bason, Wald, 2002 - reconhecimento automático da fala para geração de legendas, com taxa de erro de 80%;

- Pettit (2004) - tradução do audiovisual;
- Kurz, I. and B. Mikulasek (2004) - acessibilidade do audiovisual pelos surdos;
- Leelarammee, Mahasith,Uthaichana (2005) - mecanismo de expansão da área de LS em TVs analógicas;
- Agrafiotis et al.(2004) - codificação de vídeo em LS com redução de taxa de bits.
- Ong, Ranganath, (2005) - reconhecimento de LS;
- Papadogiorgaki, Grammalidis, Makris, (2005) - conversão de SWML para animação
- Arcoverde (2006) - importância da apropriação de português pelos surdos;
- Nakazono, Nagashima, Ichikawa (2006) - codificação de vídeo para LS;
- Wald, 2006 - correção dos erros do reconhecimento ao vivo, com possibilidade de correção de 11 erros por minuto;
- Callejas Bedregal, Dimuro, Rocha Costa (2006) reconhecimento de gestos manuais com luva de dados e saída orientada ao HamNoSys;
- Farhadi, Forsyth (2006) - alinhamento automático de legenda em LS americana com texto em inglês;
- Buttussi, Chittaro, Coppo, (2007) - busca de cheremas pela seleção de imagens de um avatar 3D;
- Caplier et al. (2007) - imagem e vídeo para pessoas surdas;
- Karpouzis et al. (2007) - tradução de texto e síntese de sinais e língua grega de sinais;
- Hooper et al. (2007) - efeitos da qualidade do vídeo na compreensão de LS ;
- Kennaway et al.(2007) - animação de avatar a partir de HamNoSys e SigML;
- Davies et al(2007) - codificação de vídeo baseada em percepção visual do surdo;
- Pistori (2007) - visão computacional aplicada a acessibilidade , reconhecimento de LS por vídeo;
- Gennari, Mascio (2007) - ontologia para busca de sinais em dicionário;

- Burnham et al. (2008) - efeitos da taxa de legendas e redução de texto na compreensão do audiovisual;
- Jemni, Elghoul (2008) – propõem uma aplicação baseada na web para criar dicionário de LS usando a abordagem wiki;
- Okura, Hirose (2008) - desenvolvem um player capaz de recuperar legendas de DVDs em uma base de dados na web;
- Kountchev, Todorov, Kountcheva, 2008 - esquema de codificação de vídeo em LS que transmite apenas os contornos da imagem;
- Xu et al. (2008) - sistema de animação de avatares em língua de sinais chinesa;
- Tania, Mauro, Rosella (2008) - esquema de busca em glossário de LS;
- Pollard et al (2009) - conteúdo para educação de surdos em LS com legendas em língua oral;
- Phan, Nguyen, Bui (2009) - agente 3D para apresentação de conteúdo digital;
- Fels et al. (2009) - multimídia para surdos;
- Adamo-Villani et al. (2009) - comparação entre tipos de avatares;
- Kamata, Fujii, Watanabe (2009) - geração de mensagens em glosa aplicado a situações de emergência;
- Shohieb et al (2009) - sistema de acessibilidade na web com geração de animação em avatar;
- Efthimiou et al. (2009)- dicta-sign trata da falta de anonimato e da facilidade de edição de contribuições online em língua de sinais;
- Cambra, Leal, Silvestre (2010) - percepção do audiovisual por surdos e ouvintes;
- Udo, Fels (2010) - design universal e acessibilidade;
- Ghoul; Jemmi, 2010 - geração de vídeo com avatar;
- Adamo-Villani et al. (2010) - Sistema de autoria de língua de sinais ;
- Möbus (2010) - web acessível aos surdos;
- Johnston, Napier (2010) - banco de sinais em língua de sinais australiana;
- Kaneko et al. (2010) - animação de avatares na TV;

- Tran et al. 2010 - aceitação de algoritmo de compactação de vídeo;
- Othman, Ghoul, Jemni, (2010);
- Ciaramello, Ko, Hemami (2010) - qualidade e inteligibilidade de vídeo em LS;
- Chon et al. (2011) - qualidade de vídeo em LS em ambientes com perda de dados.

Em relação às teses e dissertações, os resultados de Networked Digital Library of Thesis and Dissertations:

- Denardi (2006) – animação de gestos dos sinais baseada em autômatos e escrita de sinais. (dissertação)
- Teixeira (2008) – parâmetros para uma taxonomia de legendas (tese)
- Muir (2007) – codificação de vídeo aplicado em legendas de LS, baseada na percepção visual dos surdos (tese)
- Fernandes (2003) – modelo para otimização de legendas na televisão (dissertação)
- Reichert (2006) – percepção do audiovisual pelos surdos (dissertação)
- Beskow (2003) – síntese de fala com cabeças falantes. (tese)
- Lu (2010) – estratégias utilizadas na produção de legendas (dissertação)
- Maldonado (2007) – síntese audiovisual personalizável (tese)

E também os resultados de Biblioteca Digital Brasileira de Teses e Dissertações:

- Souza (2005) – os surdos e a tele informação (dissertação)
- Reichert (2006) - percepção do audiovisual pelos surdos (dissertação)
- Selvatici (2010) – análise do closed-captions (dissertação);
- Masutti (2007) - tradução entre surdos e ouvintes (tese);

- Silva(2009) – leitura de escrita de sinais (dissertação);
- Iguma(2004) - não disponível
- Zappe(2010) – uso de escrita de sinais na web (dissertação);
- Souza(2005b)- serviços na web para escrita de sinais.

Os dados obtidos demonstram que a produção científica possui grande quantidade de artigos publicados, entretanto poucas teses e dissertações diretamente relacionadas à acessibilidade do audiovisual por diferentes perfis de surdos foram encontradas.

Busca-se então identificar e analisar os fundamentos e referências necessárias para o desenvolvimento de artefatos para acessibilidade de surdos ao audiovisual e sintetizá-los/integrá-los em um modelo de referência.

ANEXO B – VALIDAÇÃO ESTATÍSTICA DAS RESPOSTAS AO QUESTIONÁRIO DELPHI

Quadro 18: Validação estatística das recomendações na primeira rodada.

Recomendação	Total de Respostas	Respostas positivas	Probabilidade	Aceitação ($\alpha < 0,01$)
1	17	17	< 0,0001	Sim
2	17	15	0,0012	Sim
3	17	13	0,0245	Não
4	17	15	0,0012	Sim
5	17	17	< 0,0001	Sim
6	16	15	0,0003	Sim
7	17	13	0,0245	Não
8	17	15	0,0012	Sim
9	15	11	0,0592	Não
10	16	8	0,5982	Não
11	15	15	< 0,0001	Sim
12	15	8	0,5000	Não
13	16	11	0,1051	Não
14	17	15	0,0012	Sim
15	17	17	< 0,0001	Sim
16	17	13	0,0245	Não
17	17	15	0,0012	Não
18	17	15	0,0012	Sim
19	17	17	< 0,0001	Sim
20	17	17	< 0,0001	Sim
21	17	15	0,0012	Sim
22	16	15	0,0003	Sim
23	17	11	0,1662	Não
24	17	17	< 0,0001	Sim
25	17	16	0,0001	Sim
26	17	17	< 0,0001	Sim
27	17	13	0,0245	Não
28	14	9	0,2120	Não
29	12	9	0,0730	Não
30	15	12	0,0176	Não
31	15	14	0,0005	Sim
32	16	12	0,0384	Não
33	15	12	0,0176	Não

34	14	9	0,2120	Não
35	15	14	0,0005	Sim

Quadro 19: Validação estatística das recomendações na segunda rodada.

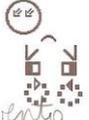
Recomendação	Total de Respostas	Respostas positivas	Probabilidade	Aceitação ($\alpha < 0,05$)
3	8	8	0,0039	Sim
7	8	7	0,0352	Sim
9	8	8	0,0039	Sim
10	8	8	0,0039	Sim
12	8	8	0,0039	Sim
13	8	8	0,0039	Sim
16	8	7	0,0352	Sim
17	8	7	0,0352	Sim
23	8	8	0,0039	Sim
27	8	8	0,0039	Sim
28	8	6	0,1445	Não
29	8	6	0,1445	Não
30	8	7	0,0352	Sim
32	8	8	0,0039	Sim
33	8	8	0,0039	Sim
34	8	8	0,0039	Sim

ANEXO C – TRANSCRIÇÕES DO CONTEÚDO EM ESCRITA DE SINAIS

Transcrição A

rãbado	defera	Minas Gias	problema	chuva
Brasil	civil	Gias	danço	chuva sim
Sul/Brasil	informar	Mato Grosso Sul	cotar/pulmão	chuva não
por causa	evitar	Mato Grosso	pode	chuva
marado	tarde	Pernambuco	São Paulo	forte
ar seco	atividade física	por um	Brasil	Pernambuco
tem	rel	ar seco	ati/dissaboleto	ati/dissaboleto
área/rua	principal	suspensão ar seco	ave	ati/dissaboleto

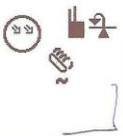
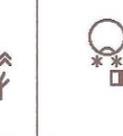
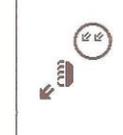
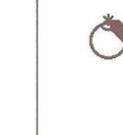
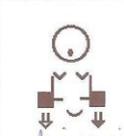
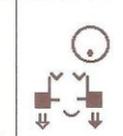
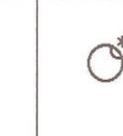
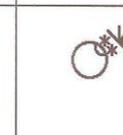
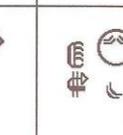
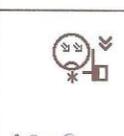
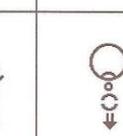
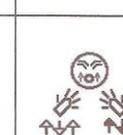
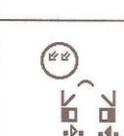
Figura 114: Transcrição do conteúdo traduzido em LS.

				
?		vento + água	Macapá	
				
graus	seco	central. região	Carutaba	
				
onze	diminuir	diminui	SP	
				
Brasília	por causa		Rio	
				
graus elevado	Brasil	aguarda	tempo	
				
33	água	tarde	hom	
				
verema	mar	calor		
				
Depois Amama	vento	chuva		

Transcrição B

 salvado	 educação educação	 Minas Gerais	 problema	 chuva
 Brasil	 civil	 Goiás	 doença	 vento
 imagem	 avisar	 Mato Grosso sul	 custar	 também
 permissão permissão	 evitar	 Mato Grosso Mato Grosso	 pode	 chuva
 chuva	 a tarde	 Rondonia	 São Paulo	 forte
 ao mesmo tempo	 esforço físico	 por causa	 Brasil	 Recife
 ter	 sol	 ao mesmo tempo	 geral	 até
 no geral	 principalmente	 doenças respiratórias	 Acre	 até principalmente do Norte

Figura 115: Transcrição do conteúdo traduzido em LS.

	 semana que sem	 mar		
 mínima	 nce			
	 diminuir			
 Brasilia	 por cause			
 máxima	 Brasil			
 ↓ ↓	 agua			
 Terzina				
	 vento			

ANEXO D – CARACTERIZAÇÃO DOS ESPECIALISTAS CONSULTADOS

O quadro seguinte descreve quantitativamente os resultados da plataforma Lattes sobre pesquisadores associados aos termos de busca.

Quadro 20: Descrição dos especialistas consultados

Formação	Termo de busca		
	“surdos + televisão”	“surdos + legendas”	“surdos+ audiovisual”
Letras	6	10	6
Educação	9	5	8
Comunicação	10	6	7
Computação	1	2	4
Artes	2	3	1
Psicologia	3	4	3
Sociologia	0	0	1

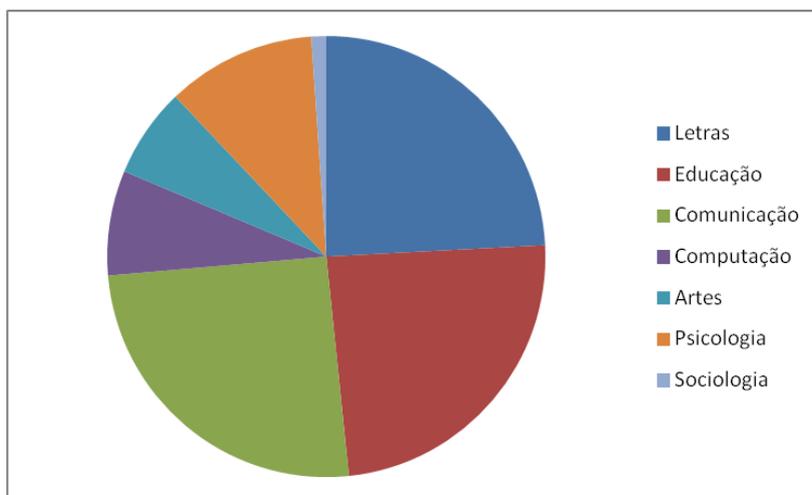


Figura 116: Caracterização dos especialistas consultados