

Edgar Correa Veras

**PROCEDIMENTOS METODOLÓGICOS PARA A COMPILAÇÃO DE UM
CORPUS DE LÍNGUA DE SINAIS A PARTIR DA REDE: REFLEXÕES COM
BASE EM UM CORPUS PILOTO DE GÊNEROS NA PLATAFORMA YOUTUBE**

Dissertação submetida ao Programa de pós graduação em linguística da Universidade Federal de Santa Catarina como parte dos requisitos necessários para a obtenção do Grau de Mestre em Linguística. Sob a orientação do Professor Dr. Tarcísio de Arantes Leite.

Florianópolis-SC
2014

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

Veras, Edgar Correa

Procedimentos Metodológicos para a compilação de um corpus de língua de sinais a partir da rede: reflexões com base em um corpus piloto de gêneros na plataforma youtube / Edgar Correa Veras ; orientador, Tarcísio de Arantes Leite - Florianópolis, SC, 2014.

183 p.

Dissertação (mestrado) - Universidade Federal de Santa Catarina, Centro de Comunicação e Expressão. Programa de Pós-Graduação em Linguística.

Inclui referências

1. Linguística. 2. Linguística das línguas de sinais. 3. Língua de sinais brasileira. 4. Corpus linguístico. I. Leite, Tarcísio de Arantes. II. Universidade Federal de Santa Catarina. Programa de Pós-Graduação em Linguística. III. Título.

Edgar Correa Veras

PROCEDIMENTOS METODOLÓGICOS PARA A COMPILAÇÃO DE UM CORPUS DE LÍNGUA DE SINAIS A PARTIR DA REDE: REFLEXÕES COM BASE EM UM CORPUS PILOTO DE GÊNEROS NA PLATAFORMA YOUTUBE.

Esta dissertação foi julgada adequada para a obtenção do Título de “Mestre em linguística” e aprovada em sua forma final pelo Programa de pós graduação em linguística do Centro de Comunicação e Expressão da Universidade Federal de Santa Catarina

Florianópolis, 25 de Março de 2014

Prof. Dr. Heronides Maurílio de Melo Moura
Coordenador do Curso

Orientador
Drº Tarcísio de Arantes Leite/UFSC

Banca Examinadora

Drª Karin Lilian Strobel/UFSC
Presidente

Drº Markus Johannes Weininger/UFSC

Drª Izete Lehmkuhl Coelho/UFSC

Drª Aline Lemos Pizzio/UFSC

AGRADECIMENTOS

Àquele que me ama suficientemente para me dar inclusive a liberdade de rejeitá-lo,

Ao meu orientador, Tarcísio Leite, que me ensinou a ser uma pessoa melhor,

À Ronice Quadros, grande responsável por toda minha trajetória acadêmica, com quem tenho orgulho de trabalhar,

À minha grande amiga de quem tive a sorte de ser filho e que testemunhou parte da escrita deste trabalho durante os quase 60 dias que passamos juntos internados,

À minha companheira inseparável Wanessa, que me deu de presente a Ana Beatriz,

Ao meu grande amigo José Ednilson, companheiro de todas as horas,

Ao meu irmão Endrigo, que primeiro me apresentou o maravilhoso mundo da surdez,

Ao meu mano Adriano, que em memória me ajudou nos momentos difíceis e ao Rodrigo, que compreendeu minha ausência,

Ao Rubens que chegou no final desta pesquisa, mas que foi fundamental para que eu pudesse concluí-la,

Aos meus colegas tradutores/intérpretes da UFSC, que me auxiliaram na conclusão deste trabalho,

À coordenação de aperfeiçoamento de pessoal de nível superior (CAPES), pelo financiamento de parte desta pesquisa e, finalmente, à toda comunidade surda, representada por aqueles presentes na pesquisa e que lutam diariamente pelo respeito aos seus direitos linguísticos e humanos.

RESUMO

O espaço em rede, através das plataformas de inserção de vídeos, tem sido cada vez mais utilizado pelas comunidades surdas como meio de interação e repositório de dados linguísticos. Constatase atualmente enorme variação e inconsistência nos critérios de registro, documentação, análise e apresentação dos dados de línguas de sinais à comunidade acadêmica, sobretudo nas pesquisas que tomam a rede como fonte. Esta pesquisa discute os procedimentos metodológicos na formação de corpus de línguas de sinais a partir da plataforma *youtube.com*, tendo a formação de um corpus de gêneros da língua de sinais brasileira (Libras) como projeto piloto para a elaboração e identificação das principais questões que cercam a constituição de um corpus a partir de vídeos da rede. Uma breve discussão sobre gêneros em corpora é apresentada a partir do levantamento proposto na pesquisa além de considerações a respeito dos procedimentos éticos adotados na pesquisa com línguas de sinais. Desenvolvemos uma análise técnico-qualitativa dos procedimentos adotados ao longo da constituição de um corpus de vídeo a partir da rede utilizando as orientações de Reppen, Koester, Thompson e Clancy (2010), Berber Sardinha (2004), Crasborn (2010, 2012), Dwyer (2006) e recorrendo a corpora de língua portuguesa e línguas de sinais já constituídos. Resultados preliminares apontam para uma possibilidade de uso da rede como uma fonte de formação de corpus desde que consideradas as adequações da finalidade de pesquisa e observados os critérios específicos para o suporte utilizado.

Palavras-chave: língua de sinais, corpus linguístico, youtube

ABSTRACT

The web space has been increasingly used by deaf communities through video uploading platforms as a means of interaction and a repository of linguistic data. Currently, one can see the presence of an enormous variation and inconsistency in the criteria for recording, documentation, analysis and submission of sign language data to the academic community, especially on researches that make the web their data source. This research discusses the methodological procedures in the establishment of sign language corpora from the youtube.com platform, having the formation of a corpus of Brazilian Sign Language (Libras) genre as a pilot project to elaborate and identify the main questions surrounding the constitution of a corpus from videos on the web. A brief discussion about genres in corpora is provided based on the data collection proposed on the research, as well as considerations about the ethical procedures adopted on the research on sign language. We developed a technical and qualitative analysis of the procedures adopted on the establishment of a video corpus based on the web, following the orientation given by Reppen, Koester, Thompson and Clancy (2010), Berber Sardinha (2004), Crasborn (2010, 2012), Dwyer (2006) and resorting to already composed corpora of Portuguese and sign languages. Preliminary results point to the possibility of using the web as a source for the establishment of corpora, once provided the adjustments for the purpose of research are considered, and the specific criteria to the chosen source of data are observed.

Keywords: sign language, linguistic corpus, youtube

SUMÁRIO

INTRODUÇÃO	11
1 FUNDAMENTAÇÃO TEÓRICA	19
2 QUESTÕES ÉTICAS NA FORMAÇÃO DO CORPUS	35
3 QUESTÕES METODOLÓGICAS PARA A FORMAÇÃO DO CORPUS	59
3.1 Da extração de vídeos na rede para a formação de corpus de línguas de sinais.....	59
3.2 Organizando o corpus de vídeo	63
3.3 Limpeza do corpus de vídeo.....	67
3.4 Seleção dos vídeos	70
3.4.1 A questão dos gêneros discursivos	70
3.4.1.1 Os gêneros em corpora	77
3.4.1.1.1 Gêneros em corpora da língua portuguesa	77
3.4.1.1.2 Gêneros em corpora de línguas de sinais	82
3.4.1.1.3 Gêneros em padrão de metadados	84
3.4.1.1.4 Gênero nas demandas da comunidade surda	86
3.4.1.1.5 Gênero na pesquisa linguística da libras	87
3.4.1.1.6 Gênero na plataforma de pesquisa.....	89
3.5 Definição dos gêneros e montagem do corpus	102
CONSIDERAÇÕES FINAIS	127
REFERÊNCIAS	131
APENDICE.....	139
ANEXOS.....	143

INTRODUÇÃO

Vivemos uma revolução no modo de compreender o mundo e organizar nosso pensamento sobre ele. A comunicação em massa, mediada majoritariamente por sistemas de escrita potencializados na invenção da imprensa, sofreu uma grande transformação com o advento de outros sistemas semióticos de comunicação em larga escala. A invenção do rádio e da televisão no século XX são fatos que autenticam essa afirmação. A invenção do telefone no século anterior, mas de acesso ampliado somente no século XX, possibilitou uma comunicação dialógica em tempo real. Esse fato permitiu a transgressão dos limites da espacialidade o que tornou os modos de comunicação ainda mais amplos e impactantes nas relações sociais.

Uma transformação peculiar surge nas comunicações com o advento das redes digitais por sua natureza radicalmente diversa dos meios de massa mais tradicionais constituindo-se plataformas de comunicação, interação e colaboração instantâneas, conforme aponta Mello (2010, p. 17), “ao permitir que qualquer indivíduo ou grupo com acesso a um computador conectado se torne, por si mesmo, um meio de comunicação, capaz de mobilizar milhões de outros, a internet estaria introduzindo um novo paradigma nas relações sociais, econômicas e políticas”.

Diversos fatos sociais podem exemplificar esse impacto e o papel que as redes digitais tomam na contemporaneidade. Em novembro de 2008, cita ainda a autora, “a eleição de um político negro no cargo mais importante do mundo deve-se principalmente ao uso da chamada web 2.0, a segunda geração da rede, que permitiu a formação dos chamados militantes on-line”. O impacto do uso da rede ficou evidente na arrecadação da campanha online que foi superior à dos candidatos rivais que utilizaram estratégias tradicionais.

Em 2010 e 2011 diversos grupos de manifestantes em países do mundo árabe, que estão há décadas sob regimes autoritários e sofrem com enormes problemas sociais, conseguiram organizar seus protestos e ir às ruas graças ao uso de plataformas da rede como youtube, twitter e facebook. (HOWARD, 2011)

A possibilidade de autoria da informação tem transformado a forma como as pessoas constroem suas posições sociais, conhecimentos e como esses conhecimentos afetam suas vidas. Toda informação agora

pode ser questionada através de fontes que protegidas pelo suposto anonimato virtual podem rebelar-se contra o monopólio da verdade.

O impacto causado nas diversas comunidades urbanas de falantes de línguas orais em todo o mundo não pode ser equiparado com aquele causado aos povos surdos de comunidades urbanas, usuários de língua de sinais. Para eles mais do que uma melhoria na forma de relacionar-se e comunicar-se esse impacto representou o rompimento da espacialidade, temporalidade e instabilidade no registro de suas línguas, de modalidade viso-gestual, nunca antes experimentado. Isso só foi possível graças ao surgimento e acesso a dispositivos de gravação de vídeo e ao aumento da capacidade de os computadores pessoais processarem e armazenarem vídeos digitalizados.

Na rede, o uso de plataformas de comunicação em vídeo, que permitem a acomodação das línguas de sinais, representa por excelência um dos espaços potencializadores dessa revolução. Crasborn (2010) refere-se ao impacto dos vídeos em rede para a comunidade surda como bem maior do que se pode prever atualmente. Tratando especificamente sobre uma dessas plataformas de rede, o site *youtube.com*, Festa (2012, p.155) aponta que “(...) através do YouTube, os surdos conseguiram estabelecer motivos para uma valorização do “ser surdo”, o que representa possibilidade de reconstrução de um novo olhar sobre a surdez por meio da língua de sinais”. Vemos aqui, no entanto, a plataforma muito mais como um elemento potencializador e facilitador da luta política do “ser surdo” que já havia sido iniciada fora desses espaços virtuais. Apesar disso, acreditamos que o impacto da plataforma youtube sobre as relações da comunidade surda, de forma possivelmente motivadora também, é significativo.

O avanço possibilitado em parte pelos dispositivos de gravação, reprodução e armazenamento de vídeo – paralelamente aos de áudio para as línguas orais – permitiu uma mudança também na ciência linguística. A possibilidade de registro e manipulação de grandes quantidades de dados linguísticos possibilitou a ênfase nas pesquisas com sistematização da observação de dados linguísticos e não mais através da indução ou idealização como tradicionalmente se realizava. (BERBER SARDINHA, 2000, 2004).

O espaço em rede emerge então na contemporaneidade como uma possível fonte de dados de produção linguística, constituindo-se base para a formação de importantes corpus de línguas orais conforme aponta Lee (2002). Citando ainda Kilgariff e Greffensette (2003, p.2)

o autor mostra que essa possibilidade emerge principalmente “a partir de uma definição abrangente de corpus que possibilita que a própria rede seja considerada um corpus em si”. Kilgarrieff (2001, p.342) destaca a importância da rede como recurso de formação de corpus de línguas orais em todo o mundo devido à “enorme quantidade de textos, variedade das amostras e facilidade de acesso”.

Podemos constatar, no entanto, que há uma grande diferença observada aqui quanto à natureza dos dados não levantada pelos autores. A rede ainda é composta majoritariamente por dados da língua escrita, o que limita na fonte a observação da língua como ela realmente se apresenta: uma produção complexa de gestos, prosódia, entonação, direção do olhar, expressões, referências e outros diversos elementos por vezes não notados pela limitação imposta pela escrita. Do contrário, o material disponível em língua de sinais está posto quase sempre em vídeo, o que permite uma análise mais completa do fenômeno linguístico.

Logo, para a formação de corpus em línguas de sinais, as plataformas de vídeo, como já mencionado, tornaram-se esse possível espaço de excelência na formação de corpus linguístico. Isso se dá principalmente devido a riqueza das situações comunicativas presente nos vídeos que circulam nessas plataformas, às variedades linguísticas disponíveis e ao alcance geográfico dos dados. Cabe mencionar ainda o aspecto da espontaneidade dos dados. Embora haja um conjunto de considerações a serem feitas quanto à influência da filmagem em vídeo na performance do sinalizador, destacamos aqui a predominância de filmagens feitas no ambiente doméstico cotidiano do produtor do vídeo, constatada no levantamento feito na plataforma *youtube.com*. Essa caracterização do ambiente produz um efeito de espontaneidade notavelmente maior que aquele obtido nos laboratórios de pesquisa.

Esta pesquisa discute, portanto, os procedimentos metodológicos na formação de um corpus de línguas de sinais a partir da plataforma *youtube.com*, tomando como referência um corpus de gêneros da língua de sinais brasileira (Libras) para a elaboração e identificação das principais questões que cercam a constituição de um corpus de vídeos.

O cenário em que esta pesquisa se instaura reflete uma trajetória recente na pesquisa linguística das línguas de sinais. Durante muitos anos, as línguas de sinais não foram consideradas línguas naturais de fato e ainda há resistências quanto a isso mesmo no meio acadêmico.

Isso se deve principalmente ao fato de a surdez estar associada a uma deficiência humana dos sentidos. Como os conhecimentos, saberes científicos e verdades de uso social são construídos a partir da ótica da maioria, a ausência de um dos sentidos tem sido constituída como verdade sob a lógica da perda, o que a surdez representa para os não surdos. Um recurso comunicativo que quase sempre resulta de uma privação, perda, prejuízo, ou deficiência (tomando sua presença em referência) não tem sido encarado em pé de igualdade com a língua “natural” desenvolvida por indivíduos falantes e “agraciados” com todos os sentidos disponíveis à raça humana.

Tem havido, portanto, uma necessidade constante que perpassa a representação social da maioria ouvinte e esforços biomédicos e pedagógicos generalizados para erradicar a perda auditiva e por consequência, as línguas de sinais, vistas como resultados patológicos dessa deficiência, conforme aponta Johnston (2004).

Os primeiros a advogarem pelo estatuto linguístico das línguas de sinais foram os próprios indivíduos surdos, membros das comunidades de surdos sinalizantes. Para isso tiveram que ser contrários à lógica majoritária que ou os impelia ao aprendizado da língua oral em proibição ao uso das línguas de sinais ou permitia seu uso apenas como recurso ao aprendizado das línguas orais. Como resultado criaram seus clubes e associações ou mantinham seus pontos de encontro e espaços “protegidos” para uso, manutenção, evolução e repasse do artefato linguístico.

Para a comunidade científica, o primeiro a defender o estatuto linguístico das línguas de sinais como equivalente ao das línguas orais, foi William Stokoe, professor e então chefe do departamento de inglês da Universidade para surdos Gallaudet, em Washington – DC, na década de 1960, com base no estudo da língua de sinais americana (ASL). Ao notar que a língua de sinais utilizada pelos alunos surdos entre si era muito mais “completa” do que a língua utilizada na interação com os professores em sala de aula, Stokoe passou a observar e descrever essa língua. (ARMSTRONG, KARCHMER e VAN CLEVE, 2002). Seus estudos o levaram a demonstrar que a ASL, e por inferência as demais línguas de sinais, possuíam as mesmas propriedades fundamentais consideradas critério linguístico para validação de uma língua e portanto poderiam ser analisadas nos mesmos níveis, utilizando-se dos mesmos procedimentos teóricos e metodológicos aplicados a essas línguas. Ao

iniciar a descrição da estruturação sublexical dos sinais, Stokoe inaugurou, então, os estudos linguísticos da língua de sinais.

O cenário em que essas pesquisas emergiram nos Estados Unidos, não ocasionalmente, era propício a uma valorização dos direitos de minorias linguísticas. Uma intensa luta pelos direitos civis de gays, negros, mulheres e dos considerados “deficientes” se instaurava no contexto norte americano no final da década de 1960. No Brasil, cenário similar se instaurou a partir do ano 2000, quando as lutas pelos direitos linguísticos dos surdos se intensificaram. Todo esse movimento culminou com a aprovação da Lei n. 10.436, de 2002, que reconhece a libras como língua da comunidade surda brasileira, e o Decreto n. 5626, de 2005 que a regulamentava.

Vários avanços puderam ser registrados desde então. Entre eles, destaca-se a obrigatoriedade do ensino de língua de sinais brasileira nos cursos superiores de licenciatura em todo o território federal, o que demandou a formação de professores e intérpretes de libras e a criação de cursos de graduação em Letras Libras por todo o país.

O estudo científico da Libras, portanto, passa hoje por uma fase decisiva. A demanda acadêmica e social por conhecimento relativo a Libras é grande, mas o campo de investigação ainda está se estruturando. Como aluno do curso letras libras da Universidade Federal de Santa Catarina e depois como monitor, tutor e intérprete na mesma Universidade, a dificuldade em lidar com os dados para pesquisa linguística da Libras ficou evidente.

Em 2010 passei a atuar no Núcleo de aquisição de língua de sinais da UFSC, coordenado pela prof. Dr^a Ronice Muller de Quadros. A questão metodológica no procedimento de formação de corpora de língua de sinais, recorrente nos projetos do núcleo, surgiram então com mais veemência. Em seguida passei a atuar como monitor do prof. Tarcísio Leite, que já vinha trabalhando com questões ligadas à formação de corpus em língua de sinais desde sua formação inicial na Universidade de São Paulo. O recorte desta pesquisa foi então definido.

A delimitação deu-se a partir da constatação de que a Libras não possui um corpus significativo da língua em uso disponível. Considerando a observação desse tipo de dados importante para a pesquisa linguística, surgiram os questionamentos: como coletar, analisar e referenciar um corpus de língua de sinais para pesquisa? Tendo a rede como uma possibilidade para a formação de corpus, de que forma os procedimentos já definidos para a formação de corpus de língua de sinais

por alguns pesquisadores que serão mencionados nesta pesquisa, se aplicam nesse tipo de trabalho?

Em razão dessa problemática, nos últimos anos, a comunidade de pesquisadores da língua de sinais em centros de pesquisa de vários países tem dedicado atenção aos procedimentos de coleta, documentação e recuperação de dados e metadados de línguas de sinais. Inúmeras propostas têm surgido para dar conta destas questões de inconsistência metodológica, mas atualmente ainda estão em fase inicial e se diferem significativamente provocando uma variação e inconsistência nos critérios de registro, documentação, análise e disponibilização dos dados de línguas de sinais à comunidade acadêmica.

Nossa proposta é que a rede pode representar uma das alternativas para as dificuldades encontradas pelos pesquisadores na captação, registro, montagem e organização de dados linguísticos para a formação de um corpus de língua de sinais. Sua importância se dá, sobretudo devido ao rico repositório de dados que ela contém, dados que dificilmente seriam obtidos de outro modo o que justificaria um novo olhar para essa fonte. Não é surpresa, no entanto, constatar também a inconsistência nos critérios de registro, documentação e análise para essa fonte em particular, considerando a escassez de trabalhos que sistematizem esses procedimentos. Acreditamos que a partir de um projeto metodológico bem delineado, corpora significativos podem ser obtidos com base na rede, estimulando um profícuo campo metodológico de pesquisa.

Logo, pretendemos aqui discutir os procedimentos metodológicos necessários para a compilação de um corpus de língua de sinais a partir de uma plataforma de vídeo na rede. Para isso propomos a compilação de um corpus de gêneros da Libras que sirva como um corpus piloto para esse processo. Com isso queremos levantar e propor reflexões quanto aos procedimentos adequados para a formação desse tipo de corpus apontando soluções e estratégias para o uso da rede como fonte de dados linguísticos de língua de sinais. Uma breve discussão sobre gêneros em corpora será feita ainda a partir do levantamento proposto na pesquisa.

Para isso inicialmente definimos os procedimentos éticos necessários para a manipulação de dados na rede com base na bibliografia geral e específica sobre o assunto. Em seguida analisamos os procedi-

mentos disponíveis para extração de dados em vídeo bem como as plataformas de hospedagem desses vídeos utilizadas pelos surdos através de um levantamento na própria rede. Com base nos resultados preliminares estabelecemos critérios para a definição das ferramentas de extração e para a escolha da plataforma de vídeo. A partir dessa definição selecionamos bibliografia específica com orientações para a formação de corpus linguístico (textos escritos) a partir da rede. Recorremos também a trabalhos que tratavam da formação de corpus de língua de sinais em laboratório estabelecendo um paralelo sempre que possível. Com base nas orientações técnicas consultadas, em parte da bibliografia disponível sobre o assunto e na concepção teórica adotada de língua e gênero, elaboramos o corpus de alguns gêneros da Libras. Finalizamos com algumas considerações sobre as adequações e inadequações dos procedimentos propostos e a busca de sistematização dos resultados, sugerindo novos padrões sempre que necessário.

Acreditamos que discutir a adequação, padronização e estabelecimento de procedimentos metodológicos para a formação de corpus da Libras irá possibilitar avanços significativos no campo de estudos da linguística de corpus de línguas de sinais, além de possibilitar a abertura de um campo que já vem sendo explorado na linguística de corpus de línguas orais: o uso da rede como fonte de dados para pesquisas com língua de sinais. Resultados preliminares apontam para uma possibilidade de uso da rede como uma fonte possível de formação de corpus desde que consideradas as adequações da finalidade de pesquisa e observados os critérios específicos para o suporte utilizado.

O trabalho está organizado da seguinte maneira: após a apresentação introdutória da pesquisa, no primeiro capítulo tratamos da fundamentação teórica que subsidia o trabalho, situando a pesquisa na perspectiva da *abordagem baseada no uso*. Definimos também nesse capítulo alguns conceitos importantes na pesquisa como o conceito de corpus, representatividade, gênero e a questão dos dados dispostos na rede, situando as abordagens utilizadas nos procedimentos de coleta, organização e reflexões sobre o processo. No segundo e terceiro capítulos realizamos a pesquisa propriamente dita que consiste nas reflexões sobre o processo metodológico de formação de um corpus da rede a partir da elaboração de um corpus piloto. Devido à importância dada às questões éticas dentre as diversas questões envolvidas nos procedimentos metodológicos, elas foram tratadas em capítulo específico (2). No capítulo 3 continuamos as discussões tratando dos demais procedimentos metodo-

lógicos adotados na formação do corpus. Por último apresentamos algumas considerações preliminares sobre os resultados do trabalho e que implicações elas parecem apontar.

1 FUNDAMENTAÇÃO TEÓRICA

Neste capítulo pretendemos apresentar as abordagens teóricas utilizadas nesta pesquisa. Iniciaremos apresentando a concepção geral de língua a que essa pesquisa se filia para depois situá-la mais especificamente no campo da linguística de corpus. Trataremos em seguida da relevância dos dados dispostos na rede para depois definirmos as abordagens utilizadas nos procedimentos de coleta, organização e reflexões sobre o processo bem como a questão da representatividade e gênero.

Partindo da concepção de que as teorias determinam a forma como o pesquisador apreende o objeto de estudo e os resultados obtidos, consideramos importante definir a teoria linguística que tomamos como base aqui. Logo, os procedimentos metodológicos para a formação de um corpus linguístico estão diretamente relacionados com a forma como a linguagem é concebida pela abordagem teórica que a sustenta.

Martelotta (2010), Bentes e Mussalim (2004) e Lyons (1987) discorrem sobre as diversas abordagens teóricas que tem tratado do objeto da ciência linguística. A abordagem estruturalista, que remonta ao início da ciência linguística do séc. XX contempla a língua enquanto forma, organizada em um sistema, regida por um conjunto de leis internas. A abordagem gerativa, que veio logo após, continua considerando a língua enquanto um sistema homogêneo regido por regras mas acredita que esse sistema é inato e apenas acionado no contato com o social. Diferente dessa perspectiva formalista da língua, uma abordagem que relaciona a língua com a sociedade, a funcionalista, passa a focar seus estudos na função das unidades linguísticas e como elas influenciam a formação do sistema linguístico. A língua passa a ser vista de forma heterogênea e observada nas diferentes situações comunicativas. Seguindo nessa abordagem da língua enquanto elemento dinâmico que está em constante mudança através dos seus mecanismos de codificação emergentes, a abordagem da sociolinguística foca seus estudos nos processos de variação e mudança, tentando demonstrar a sistemática covariação entre a estrutura linguística e social. Dentro desse contexto surge uma nova visão integradora da linguagem, que alia não só linguagem e sociedade, mas cognição. A língua não estaria compartimentalizada na mente humana, mas seria regida por uma estrutura conceptual mais am-

pla que é usada em outros âmbitos da experiência humana o que constitui o cerne da abordagem cognitiva.

De forma sucinta, as abordagens que compartilham de uma visão de que a gramática não é apenas um sistema de produção e compreensão da linguagem, mas também é formada por esses processos durante a interação linguística são chamadas de abordagens baseadas no uso. Newmeyer (2003) destaca o surgimento do termo citando o trabalho de Langacker (1987: 494) que “*cunhou o termo "modelo baseado em uso" para se referir a essas abordagens que rejeitam uma nítida distinção entre o conhecimento e o uso da língua.*” Embora o contexto e o trabalho de Langacker sugere que ele estivesse se referindo a um âmbito muito menor da influência do uso na linguagem (TORRENT, 2012), o termo abrange hoje uma ampla gama de questões e métodos bem maiores onde o entendimento de linguagem adotado nesta pesquisa se situa.

Considerando a existência de vários pontos de vistas diferenciados sobre a abordagem baseada no uso, o ponto chave aqui é o reconhecimento da importância da experiência do usuário com a linguagem, o uso. Essa experiência iria interagir com o aparato cognitivo do usuário para formar as representações mentais que tornam o uso da linguagem ainda mais produtivo. (BYBEE, 2010) Logo, a formação de corpora eletrônicos que permitam a sistematização e a manipulação dos dados linguísticos de uso tem sido considerada fundamental por grande parte dos linguistas cognitivos, funcionais e sociolinguistas sob o escopo dessa abordagem. Essa perspectiva mais ampla e central é resumida por Ibbotson (2013, p. 1):

Despite the daunting scope of linguistic phenomena begging an explanation, usage-based theories of language representation have a simple overarching approach. Whether the focus is on language processing, acquisition, or change, knowledge of a language is based in knowledge of actual usage and generalizations made over usage events. (Langacker, 1987, 1991; Croft, 1991; Givón, 1995; Tomasello, 2003; Goldberg, 2006; Bybee, 2010)

A ênfase em uma pesquisa focada na constituição de corpus e nos procedimentos metodológicos subjacentes se dá, principalmente, por

entendermos que a pesquisa que toma por base a língua em uso pode evidenciar questões importantíssimas sobre sua natureza. Se é na interação que a gramática se constitui e os padrões de uso exercem influência sobre ela que por sua vez imprime restrições sobre as ações com a linguagem, a observação do uso nos trará valiosas informações sobre sua constituição. Conforme Bybee (2010, p. 1):

(...)um dos principais objetivos da abordagem baseada no uso é explorar a possibilidade de que os fenômenos estruturais que nós observamos na gramática das línguas naturais podem ser derivados dos processos cognitivos de domínio mais gerais e da forma como eles atuam em várias instâncias do uso da língua”.

Embora trate apenas dos procedimentos metodológicos que levam a esse tipo de reflexão e análise, entender que partimos dessa abordagem é fundamental também para o processo do tipo inicial de formação do corpus, tal qual alvo de reflexão aqui, pois delimita também as escolhas, a disposição dos dados na pesquisa, e toda a estrutura prévia de disposição do corpus. Portanto, a proposta que assumimos aqui é uma tentativa de compilação de um corpus que possibilite uma análise translinguística da forma e função (com destaque para a última) das práticas discursivas e todos os elementos envolvidos.

Situado o panorama geral da abordagem linguística base da pesquisa, delimitemos um pouco nosso campo de estudos. Quando falamos de corpus e formação de corpora da língua de sinais, nos situamos dentro de um campo que tem se desenvolvido bastante nos últimos anos, o da linguística de corpus (LC). É importante, contudo, destacarmos alguns recortes e definições que o campo tem estabelecido e, com isso, entendermos as concepções utilizadas neste estudo. Segundo Berber Sardinha (2004, p.18) um corpus pode melhor ser definido como:

Um conjunto de dados linguísticos (pertencentes ao uso oral ou escrito da língua, ou a ambos), sistematizados segundo determinados critérios, suficientemente extensos em amplitude e profundidade, de maneira que sejam representativos da totalidade do uso linguístico *ou de algum de seus âm-*

bitos, dispostos de tal modo que possam ser processados por computador, com a finalidade de propiciar resultados vários e úteis para a descrição e análise. (*Grifo nosso*)

Essa definição é mais adequada para o autor porque menciona vários pontos importantes, como a origem, o propósito, a composição, a formatação, a representatividade e a extensão. Consideramos que essa visão se adequa àquela estabelecida aqui, principalmente por atribuir ao corpus linguístico a capacidade de ser representativo *de alguns dos âmbitos da língua*, o que propomos para um corpus de gêneros em uma plataforma de vídeo.

É importante destacar, no entanto, que o autor parece limitar o trabalho da linguística de corpus a uma visão probabilística da linguagem. Isso implicaria resumir corpus a dados que pudessem ser processados e analisados por computador. Dessa forma as ferramentas centrais seriam as computacionais de análise, como os etiquetadores, listagem de palavras, concordanciadores, entre outros. O trabalho do linguista teria como foco, nesse prisma, a investigação da frequência de ocorrência de traços linguísticos (lexicais, sintáticos, semânticos, discursivos) a partir de análises estatísticas em corpus com grandes quantidades de dados.

O estado atual de desenvolvimento e disponibilização de ferramentas computacionais de análise para uso amplo e livre em corpus de línguas de sinais impossibilita que nos enquadremos em um tipo de trabalho restrito a busca de padrões linguísticos através de resultados estatísticos. Uma exceção seriam os trabalhos que analisam os efeitos de frequência lexical na variação e mudança fonológica e morfológica da língua de sinais a partir do corpus da língua de sinais Australiana, Neozelandesa e Britânica (SCHEMBRI e JOHNSTON, 2011). Ainda assim, reconhecendo a necessidade de corpus mais robustos para a avaliação de frequência, os autores mencionam:

Although larger, more representative corpora and greater use of more consistent lemmatisation procedures are both needed before more robust measures of lexical frequency in sign languages can be made available, these results do at least lay the groundwork for an understanding of the rela-

tionship between token frequency and usage-based models of sign language grammar. (pág. 10)

Segundo Berber Sardinha (2004), a representatividade estaria ligada inicialmente à extensão, ou seja, quanto maior o corpus mais representativo ele seria. Essa visão, no entanto, está relacionada à perspectiva do autor de que o levantamento da probabilidade de ocorrência de traços lexicais, estruturais, pragmáticos e discursivos está no cerne da linguística de corpus (p. 24). Entendemos como já dito, ainda não ser possível realizar trabalhos com relativa extensão numérica devido às limitações de ferramentas de trabalho para as línguas de sinais.

O autor menciona, contudo, alguns contra-argumentos que tornam a questão da representatividade complexa, conforme resumimos aqui: a amostragem representativa só poderia ser determinada a partir de uma população que pudesse ser dimensionada com exatidão, o que não é o caso de usuários de uma língua (p. 23); a restrição de categoria (como limitar o corpus a um autor, gênero ou outra categoria) possibilitaria atingir um maior grau de fechamento e uma melhor representatividade (McENERY e WILSON, 2001); corpora compilados em pequena escala por pesquisadores individuais acabam sendo mais representativos do que os respectivos subcorpora extraídos de grandes corpora gerais (pág. 28); a adequação de um corpus não é medida somente pelo tamanho mas pelo que se busca na pesquisa (pág. 29).

Corroborando com os argumentos de Berber Sardinha, acreditamos que a proposição de um corpus de alguns gêneros da língua de sinais dispostos na rede, então, cumpre sua função de representatividade de alguns âmbitos da língua à medida que se restringe a um suporte único e, portanto, se limita a representar a variedade da língua relacionada à natureza do suporte apresentado. Por não possuir ainda uma estrita finalidade de pesquisa de aspecto fonológico, morfológica, sintático ou de algum fenômeno linguístico em particular e não ter finalidade de levantamento de formas *hapax legomena* ou de alta frequência não está sujeito a obrigatoriedade de extensão. Entendendo a padronização da linguagem como fenômeno relativamente estável em situações comunicativas semelhantes, as amostras, se cuidadosamente selecionadas, mesmo em menor número, podem refletir a realidade proporcional em corpora maiores. Tratando da questão do tamanho da amostra e sua representatividade, Biber (2008, p. 63) menciona:

Typically researchers focus on sample size as the most important consideration in achieving representativeness: how many texts must be included in the corpus, and how many words per text sample. Books on sampling theory, however, emphasize that sample size is not the most important consideration in selecting a representative sample; rather, a thorough definition of the target population and decisions concerning the method of sampling are prior considerations.

Serviriam esses dados, portanto, no mínimo para indicação de padrão e material de verificação em futuros trabalhos quantitativamente mais amplos. Por fim, observamos ainda, que as complexas atividades de domínios cognitivos que podem ser depreendidas do fenômeno linguístico não dependem estritamente de quantidade de dados, mas de análise cuidadosa dos mesmos.

Propomos, então, uma pesquisa focada no uso pela riqueza dos dados linguísticos desse tipo e acreditamos que uma amostragem criteriosa e representativa, mesmo limitada em quantidade, pode indicar elementos importantes da padronização da linguagem, bem como possibilitar generalizações e reflexões teóricas que possam promover *insights* e avanços na área. Considerando que há uma correlação entre características linguísticas e situacionais, o estudo mostra-se relevante para indicar padrões específicos dos contextos de gêneros definidos na plataforma e possibilitar futuras análises de regularidades entre os mesmos gêneros em suportes diferentes e mais amplos.

Tomamos como ponto de partida a importância que a rede assume e sua riqueza como espaço privilegiado de circulação de dados linguísticos. Isso se dá principalmente pela quantidade e variedade desses dados que possibilitam uma coleta expressiva e fornecem material para investigações linguísticas de diversos tipos. Dentre as plataformas de vídeo disponíveis, se destaca a plataforma *youtube.com*. Diversos fatores colocam esse canal em destaque como a centralidade do receptor no processo de comunicação (BERTI, 2012), o pioneirismo nos processos de digitalização de vídeos e interatividade (SERRANO, 2009), a popularidade, o uso tanto por veículos tradicionais de comunicação quanto por usuários comuns (FURUNO, 2010) e por ter se inserido no cotidiano das pessoas integrando uma política de cultura popular participativa (BURGUESS; GREEN, 2009). Além destes, é a plataforma de

vídeo que possui a maior quantidade de vídeos (+ de 460 milhões) e a maior quantidade de acessos por dia (1 bilhão e 200 milhões). É apresentado por diversos sites de ranking da categoria como o melhor site de hospedagem e compartilhamento de vídeos e segundo o site ALEXA¹, especializado em contagem de acessos na rede, o youtube é o terceiro site mais visitado do Mundo.

Dentre estes diversos fatores, um deles tem particular importância para a pesquisa linguística; a possibilidade de inserir-se no cotidiano das pessoas e com isso servir como repositório de dados que retrata diversas variedades de uso linguístico. De fato, conforme vemos no prefácio da obra de Burgess e Green (Ibid), pesquisadores australianos de inovações tecnológicas, a plataforma já é vista como ferramenta integrada ao dia-a-dia dos usuários da rede. Em trabalho na rede que resume a obra referida, Almeida (2010), menciona:

Não seria demais dizer que o YouTube consiste em uma esfera cultural pública, potencializado a cidadania cultural cosmopolita, uma vez que os vídeos mais frequentes armazenados no site dizem respeito à vida cotidiana dos cidadãos, seus valores, pensamentos e cultura

Para a comunidade surda usuária de língua de sinais, mais do que apenas uma nova possibilidade de comunicação a plataforma representou uma possibilidade de utilização de suas línguas, de modalidade gestual-visual em uma perspectiva já experimentada por falantes de línguas orais, a do rompimento do tempo e do espaço.

Alguns pesquisadores têm refletido sobre o impacto das plataformas em vídeos na rede para as comunidades surdas. Crasborn (2010, p. 279) em excerto sobre os avanços tecnológicos e os estudos das línguas de sinais, menciona:

Moreover, it has become common to view and share video files online. While the currently most popular web site, YouTube, may well disappear within a few years, the new concept of publishing

¹ disponível em <http://www.alexa.com/siteinfo/youtube.com>

will not. (...) From the perspective of Deaf communities, there is a different prominent aspect to these developments.(...) the online publication of collections of sign language video material may have a greater impact on Deaf communities than may be foreseen.

Em trabalho posterior, Hiddinga e Crasborn (2011), discorrendo sobre o uso de sinais internacionais pela comunidade surda, apontam a enorme popularidade que o youtube ocupa nessa comunidade e sugerem que os vídeos do youtube desempenham o papel que a escrita representa em comunidades de falantes de línguas orais. Apontam, ainda, os autores a capacidade da plataforma de estimular uma padronização no uso de sinais internacionais pelas comunidades surdas de diversos países e um intercâmbio maior entre elas. O mesmo aponta o pesquisador surdo Schallenberger (2010, pág. 12):

Em comparação a outras formas de expressão utilizadas pelos surdos (desenhos, charges, teatro, mímica, gestos, metáforas), o youtube parece se destacar e ganhar espaço como uma ferramenta que abre outras possibilidades, pois parece abranger diferentes formas de expressão. Sendo um meio de registro inimaginável há tempos atrás por sua agilidade, acredito que traz uma nova maneira de pensar a circulação da língua de sinais, partindo de meio de expressão efêmero e dependente de tecnologias descritivas (e eventualmente filmagem com câmeras) para um registro efetivo e coletivo, semelhante ao modo como as línguas faladas possuem a escrita como apoio para registrar e guardar memórias.

No Brasil, recentemente um movimento político em luta por uma educação bilíngue para surdos e contrário à política educacional vigente foi organizado através do youtube e facebook. Analisando a importância dessas plataformas de rede para o sucesso desse movimento, Silveira e Amaral(2012) concluem:

A Internet, o YouTube e o Facebook também são um meio de transformação e luta social. O blog

Movimento Surdo em Favor da Educação e da Cultura é um dos meios encontrados por grupos que tinham uma causa negligenciada por muitas pessoas e que hoje têm a oportunidade e abertura de participarem do debate sobre o tema no país. Através do ciberativismo, o movimento surdo ganha um destaque para que suas causas não sejam esquecidas e, cada vez mais, respeitadas.

A Revista nº 44 da Federação Nacional de Educação e Integração de Surdos² aponta o vídeo publicado no youtube pelo professor, pesquisador, ator e ativista político surdo brasileiro Nelson Pimenta, como o estopim do início de todo o movimento que culminou com duas grandes passeatas na capital federal, seminários e passeatas em quase todas as capitais brasileiras e mudança na política educacional do Ministério da Educação. A seguir podemos visualizar imagem publicada na revista que atribui ao vídeo publicado no youtube a origem da mobilização. No lado direito, podemos ver imagem com a visualização do vídeo referido na plataforma youtube, com quase 10 mil visualizações:

² Entidade filantrópica, sem fins lucrativos com finalidade sócio-cultural, assistencial e educacional que tem por objetivo a defesa e a luta dos direitos da comunidade surda brasileira. É hoje a maior entidade representativa dos surdos brasileiros.

Figura 1 – Foto do educador surdo Nelson Pimenta em mobilização política na capital brasileira e visualização do vídeo publicado pelo mesmo em protesto aos rumores de possível fechamento do Instituto Nacional de Educação dos Surdos - RJ



Fonte: foto da esquerda: Edição n 44 da Revista da Federação Nacional de Educação e Integração dos Surdos; foto da direita: visualização do vídeo postado pelo educador surdo Nelson Pimenta em resposta aos rumores de fechamento do Instituto Nacional de Educação dos Surdos disponível em <<https://www.youtube.com/watch?v=bs4wzygcsq>>

Embora não possamos mensurar exatamente a dimensão do impacto dos vídeos na comunidade surda, não há dúvidas quanto à relevância dos dados e da plataforma para a pesquisa linguística.

Desenvolvemos uma análise técnico-qualitativa dos procedimentos adotados ao longo da constituição de um corpus de vídeo a partir da rede. Para isso utilizamos como base as orientações gerais para formação de corpus de Reppen(2010), Koester(2010), Thompson(2010) e Clancy (2010) e Berber Sardinha (2004), as orientações específicas para compilação a partir da rede de Berber Sardinha (2004) e Crasborn (2010, 2012) que propõe orientações para procedimentos específicos com corpus de línguas sinalizadas. Utilizamos também como referência o trabalho de Dwyer (2006) e as diretrizes éticas da Sociedade de linguística da América³ (SLA) que orientam quanto aos procedimentos éticos na pesquisa linguística, Hochgesang et al (2010) além de Crasborn (2010) e as diretrizes éticas da Sociedade de Linguística da língua

³ disponível em http://www.linguisticsociety.org/files/Ethics_Statement.pdf

de Sinais⁴ (SLLS) que orientam quanto aos procedimentos éticos para a formação de corpus de língua de sinais. Outra referência importante foram os procedimentos adotados nos corpora da língua portuguesa de Davies, Mark e Michael Ferreira⁵ (2006), AC/DC - O Corpus Brasileiro⁶ e nos corpora de língua de sinais que buscavam ser representativos da língua de sinais pesquisada retirados do levantamento geral de corpus de língua de sinais feito por Konrad (2011).

À medida que as orientações e procedimentos disponíveis nas literaturas selecionadas iam sendo avaliadas e/ou aplicadas na elaboração do corpus piloto, paralelos eram estabelecidos com os corpora de referência e reflexões eram propostas sobre as possibilidades de adequações e convergências finalizando com propostas de sistematização.

Para se chegar à definição da plataforma, analisamos as plataformas de vídeos disponíveis na rede em um levantamento feito das principais delas, considerando o critério de quantidade de vídeos postados e resultados para a busca *língua de sinais/sign language*. Definida a plataforma youtube, precisamos definir que tipo de corpus seria compilado a partir dela. Propomos a realização de um corpus representativo que abrangesse os principais gêneros disponíveis na plataforma. Consideramos gênero aqui de forma ampla como “tipos relativamente estáveis do discurso” (BAKHTIN, 1997, pág. 279) e como tais estão intimamente ligados à nossa situação cotidiana, existindo como mecanismos de organização das atividades sociocomunicativas do dia a dia (SCHNEUWLY, 2004). Logo, se definem por aspectos sociocomunicativos e funcionais determinadas pelo canal, estilo, conteúdo, composição e função que vão produzir determinada estabilidade estrutural (MARCUSCHI, 2010). Essa conceituação de gêneros será melhor trabalhada durante a pesquisa quando os procedimentos iniciais da formação do corpus estiverem sendo discutidos.

Para a definição dos gêneros a serem compilados a partir da plataforma propomos um cruzamento de dados que chegue a uma matriz dos principais gêneros dispostos em: protocolo de padrão de metadados

⁴ disponível em <http://www.slls.eu/index2.php5>

⁵ disponível em <http://www.corpusdoportugues.org/>

⁶ disponível em <http://www.linguateca.pt/acesso/corpus.php?corpus=CBRAS>

- ISLE Meta Data Initiative (IMDI)⁷; no corpus da língua portuguesa - Davies e Ferreira(2006) e Corpus AC/DC; Corpus representativos de línguas de sinais – a partir do levantamento feito por Konrad (2001); levantamento com acadêmicos surdos através de questionário e levantamento de amostragem de vídeos postados na plataforma que serviu como fonte. A Partir da matriz estabeleceram-se sete principais gêneros prévios a serem pesquisados na plataforma.

A dificuldade em tomar uma plataforma de vídeo como objeto de estudo é descrita por Burgess e Green (pág. 23;24) principalmente por marcas como a instabilidade na organização dos dados, interface, diversidade de conteúdo, frequência cotidiana análoga e multiplicidade de funções, apresentando desafios epistemológicos e metodológicos para as áreas de ciências humanas e sociais.

Investigamos algumas metodologias propostas para pesquisas na plataforma youtube de forma a contribuir com os procedimentos para o levantamento realizado aqui. Resumimos abaixo algumas dessas pesquisas.

Em Burgees e Green (Ibid) os autores fazem um levantamento de 4.320 vídeos classificados pelo site como mais populares entre agosto e novembro de 2007, buscando padrões no material da amostra e apurando fragmentos de textos individuais usando métodos qualitativos. (pág.27). De forma parecida, Festa (2011), desenvolve pesquisa qualitativa, de abordagem sócio-histórica analisando discursos de surdos sobre a surdez em vídeos postados na plataforma. Realizando uma busca a partir da palavra-chave surdez a autora recolheu uma amostragem de 759 vídeos. Desses, pré-selecionou 200 vídeos classificados como “mais visualizados”. Após assistir os vídeos, 25 se enquadraram no rótulo “textos apresentados por indivíduos surdos sobre questões relacionadas a surdez”. Considerando o critério de objetivo de cada vídeo o autor organizou em três categorias que emergiram da própria amostra e procedeu com a análise. Serrano (2010), tentando averiguar a capacidade de elevação cognitiva através da experiência audio-visual do youtube, utilizou a categorização disponibilizada na própria plataforma e fez um levantamento inicial de 1120 vídeos e 983 canais. Após assistir todos os vídeos desenvolveu uma proposta de categorização dos vídeos a partir de um agrupamento conceitual e não apenas da forma disposta no site. Após a readequação nas novas categorias, organizou os vídeos por gra-

⁷ disponível em <http://www.mpi.nl/imdi/>

dação de adequação a categoria encontrada analisando uma amostra dos protótipos de cada categoria. Por fim, o pesquisador surdo, Schallenberg (2010) buscando verificar as produções culturais humorísticas dos surdos na plataforma youtube, utilizou o sistema de busca da própria plataforma através de duas palavras chaves sugeridas pela experiência de contato do pesquisador com a segunda língua como as que melhor representavam uma titularidade possível para os vídeos inseridos pelos usuários surdos. As palavras foram “humor surdo” e “piada surdos” em português e em inglês.

Considerando os objetivos da pesquisa linguística e os sete gêneros previamente definidos a partir da matriz de dados, utilizamos o sistema de buscas por palavras chaves da própria plataforma para obter uma primeira amostragem de vídeos supostamente relacionados ao gênero. Para o universo de resultados de cada uma das sete buscas, fizemos um recorte prévio dos 100 primeiros resultados para cada gênero. Uma análise qualitativa de cada gênero permitiu fazer um levantamento inicial de informações sobre o vídeo e uma pré-seleção do material que resultou em números diferenciados de resultados coerentes com o proposto inicialmente. Após a pré-seleção dos vídeos chegou-se aos autores dos mesmos. As demais informações de metadados foram completadas configurando-se o corpus final da pesquisa. Tomou-se o cuidado, no entanto de considerarmos as particularidades das práticas sociais das comunidades surdas e como elas se materializam na configuração de gêneros. Para isso, assistimos cada vídeo da amostragem, de forma a perceber a partir das informações prévias dos metadados e das informações linguísticas apresentadas qual configuração de gênero poderia emergir no vídeo. Este procedimento possibilitou a formação de novas reconfigurações de gêneros a partir de similaridades quanto ao estilo, conteúdo, composição e função nos vídeos da amostra selecionada. Isso porque entendemos conforme Messineo (2008) que discute a documentação de gêneros da fala em comunidades indígenas do Gran Chaco na Argentina, que podem haver configurações específicas do que aquelas tradicionais dispostas em corpus de línguas orais que possuem escrita nas análises de línguas ágrafas:

In attempts to systematically study the discourse of the indigenous languages of the Americas, several works (Bauman 1975; Hymes 1981; Sherzer 1982, 1987; Woodbury 1985; Briggs 1988; Bauman and Briggs 1990; and Sherzer and Urban

1986, among others) have argued that specific indigenous speech genres are based on specific discursive units and so constitute, in this respect, forms of poetry or verbal art. These units—marked by prosodic features, intonation, and other formal characteristics such as discursive connectors, parallelism, and repetition—structure the discourse at many levels and allow genres and specific styles to be defined in languages with an oral tradition.

Concluimos este capítulo delimitando o cenário ao qual este trabalho faz parte: o de pesquisas recentes com língua de sinais instaurado na Universidade Federal de Santa Catarina. Constitui-se em um espaço conquistado por pesquisadores surdos e aliados ouvintes que tem permitido a construção científica a partir da perspectiva dos indivíduos surdos e sua comunidade. De fato, são suas “mãos” que estão presentes neste trabalho que tem como objetivo mais amplo o avanço no reconhecimento social da língua de sinais e a conquista do respeito aos direitos humanos linguísticos dos surdos. Por servir aos pesquisadores surdos principalmente, definidos como pessoas visuais (BAHAN, 1989b apud WILCOX & WILCOX, 2005), buscamos, sempre que possível, fazer referências visuais aos procedimentos adotados onde tradicionalmente haveria descrições em português, daí a abundância de imagens. Tratando-se de uma plataforma virtual, onde os textos se constroem quase sempre em hipertextos, com imagens, ícones e outros recursos, tentar capturar ao máximo essa visualidade pareceu mais apropriado.

A pesquisa se inscreve também em uma linha metodológica inaugurada recentemente no campo dos estudos da língua brasileira de sinais que traz a abordagem metodológica dos dados espontâneos de produção para a centralidade, em oposição a um espaço que foi tradicionalmente ocupado pelas abordagens clássicas estruturalistas. Com isso filia-se a um campo teórico denominado abordagem baseada no uso que percebe a fala/performance não como elemento caótico e portanto não passível de análise, mas vê, nos dados reais de uso da língua e em sua descrição, os principais elementos para a elaboração de uma teoria linguística na qual a gramática emerge na interação, no contexto de uso.

Tomamos como posicionamento crítico portanto a concepção de que embora a relativa estabilidade encontrada na gramática das lín-

guas, dentre elas a da língua de sinais, possa ser depreendida de trabalhos com foco em sua estrutura, a ligação estabelecida com o uso da língua é tão estreita que impede uma análise que se depreenda deste último, tendo a compreensão do fenômeno linguístico dependência fundamental dos processos de uso das línguas somente observados através de formação de corpus de produção espontânea ou semi-espontânea da língua.

2 QUESTÕES ÉTICAS NA FORMAÇÃO DO CORPUS

No capítulo 2 e 3 trataremos da parte central de aplicação da pesquisa. A partir das orientações disponíveis em bibliografia selecionada, pesquisas que utilizam a rede como fonte para formação de corpus, dos projetos atuais de desenvolvimento de corpora de línguas de sinais e de um recorte e análise de gêneros em corpora proposto nesta pesquisa, vamos refletindo sobre os aspectos técnicos e metodológicos que surgiram durante o planejamento e compilação do corpus e propondo a adequação de procedimentos a partir da explanação daqueles adotados aqui. Devido a amplitude e importância das questões éticas tratadas, dedicamos um capítulo exclusivo para essas considerações.

As discussões mais recentes sobre ética na pesquisa linguística, tomam como ponto de partida a relação estabelecida entre o pesquisador e a comunidade pesquisada. A perspectiva que iremos adotar aqui é aquela defendida por Dwyer (2006, pág. 32) que vai de uma pesquisa sobre a comunidade, para um tipo de pesquisa *sobre, para e com* a comunidade, que no caso deste trabalho, refere-se a comunidade surda⁸. O trabalho de Dwyer (2006) é considerado um dos trabalhos chave pelo pioneirismo no aprofundamento das questões relacionadas à ética em pesquisa linguística. Seu trabalho estabelece cinco princípios éticos fundamentais na pesquisa com documentação linguística, que traduzimos e discutimos aqui:

1 – Não fazer mal aos informantes/comunidade pesquisada incluindo prejuízos não intencionais. Inicialmente isso exige avaliar a posição do informante perante a sua comunidade. Ela o considera alguém adequado para dar informações sobre a língua? Caberia uma avaliação cuidadosa do que seria considerado “prejuízo” à comunidade envolvida. Essa avaliação inclui considerar em que medida a exposição de informações dadas por um indivíduo poderia prejudicá-lo, desacreditá-lo ou constranger o mesmo ou a comunidade envolvida. Isso pode ocorrer por exemplo, se essas informações são considerados “segredos”

⁸ Embora usemos o termo comunidade, o conceito adotado aqui é aquele defendido por Lane, Hoffmeister e Bahan (1996) apud ⁹Magnani (2007) que preferem a expressão *mundo surdo* para se referir a uma comunidade formada por indivíduos que usam a língua de sinais e se identificam com a cultura surda.

culturais pela comunidade ou permitem leituras que possam inferiorizá-la perante outra comunidade envolvida na pesquisa. A exposição do nome do informante deve ser avaliada, pois pode causar danos à privacidade tanto pela sua exposição quanto pela omissão ao não dar os créditos devidos. Alguma forma de compensação ao trabalho dada ao informante também poderia estimular a estigmatização deste. Uma compensação adequada, onde não há proventos financeiros, poderia ser pensada a partir de propostas que ofereçam o mesmo conhecimento linguístico da língua oral em relação aquele obtido da língua de sinais em questão através dos dados coletados.

2 – Reciprocidade e equidade que envolveria uma relação consultiva, continuamente negociada e respeitosa entre as comunidades. Isso envolveria também uma renegociação de metodologias e objetivos durante o processo. O princípio é não estabelecer um ponto de vista único, mas respeitar os diferentes pontos de vista das comunidades envolvidas. Um aspecto importante sobre compensação é a constatação de que quase sempre, na pesquisa linguística, os benefícios são maiores para o pesquisador do que para a comunidade, o que faz com que os pesquisadores repensem a questão da equidade nos benefícios da pesquisa;

3 – Fazer algo bom tanto para a ciência quanto para a comunidade que consiste em não apenas pensar em algo a oferecer para o progresso da ciência mas também a oferecer para a comunidade, além da compensação ao informante individual. Em geral, isso envolve algo de valor imaterial como a produção de um livro, filme ou documentário;

4 - Obter o consentimento informado antes de iniciar a pesquisa que consiste em estabelecer todos os termos previstos do acordo de pesquisa em um documento a ser assinado pelos participantes antes do início da pesquisa. A autora reconhece que, em alguns contextos, os contratos verbais são mais propícios que um contrato escrito mas atenta para o fato de que eles não cumprem os requisitos legais gerais exigidos pelo meio.

5 - Arquivar e divulgar seus dados e resultados que consiste em não permitir que os dados sejam esquecidos e guardados nos laboratórios de pesquisa inacessíveis e em pior hipótese que desapareçam à medida que os núcleos, laboratórios de pesquisa se desfaçam ou os pesquisadores não deem continuidade ao trabalho. Uma das exigências do trabalho de campo atualmente, portanto, é o arquivamento e difusão dos dados.

A autora estabelece de maneira geral como os princípios éticos devem ser pensados em cada fase de pesquisa. Elaboramos um quadro com a síntese dessas orientações.

Fase	Orientações
Planejamento	Identificar todos os potenciais participantes incluindo instituições patrocinadoras e fazer uma estimativa de remuneração para participantes locais
Ida à campo	Estabelecer e manter contatos, negociar contratos e protocolo para a obtenção dos dados
Análise	Decidir sobre o número de níveis minimamente adequados de anotações. O tipo de anotação que será feito irá determinar o público específico que irá acessar o material
Arquivamento	Atender o desejo de informantes quanto ao anonimato. Decisões devem ser tomadas sobre o acesso de usuários (comunidade, pesquisadores, público em geral) e que materiais serão acessados
Produto final	Colaborar com a elaboração de manual de diretrizes éticas para a região específica em que os dados forem coletados

A autora discorre ainda sobre *normative ethics*, (p. 33) que seriam diretrizes éticas definidas por pesquisadores do campo de estudos.

De fato, a necessidade de a comunidade de linguistas que pesquisa línguas de sinais sistematizar as diretrizes éticas desse tipo de pesquisa vêm sendo apontada há algum tempo por linguistas surdos. Conforme citado em Hochgesang et al (2010):

The Amsterdam Manifesto, prepared by a group of sign linguists following the meeting of the conference on Theoretical Issues in Sign Language Research in Amsterdam in 2000, raises the point that much of sign language research is dependent on Deaf research assistants as well as data from Deaf native signers. The manifesto suggests that one way to acknowledge the contributions from these sign language communities is to give something back to them

Os autores apontam ainda que o primeiro trabalho realizado nesse sentido foi feito por Harris, Holmes & Mertens (2009) que sistematizaram diretrizes, citadas pelos autores e traduzidas abaixo (tradução minha):

1. A autoridade para a construção de significados e conhecimentos dentro da comunidade de língua de sinais cabe aos membros da comunidade;
2. Pesquisadores devem reconhecer que os membros da comunidade de língua de sinais têm o direito de ter esses princípios que eles valorizam plenamente considerados em todas as interações;
3. Pesquisadores devem levar em conta as visões de mundo da comunidade de língua de sinais em todas as negociações ou relações que tenham impacto sobre membros da comunidade;
4. Na aplicação de termos de referência das comunidades de língua de sinais, os investigadores devem reconhecer as experiências diversas, entendimentos, e modo de vida (nas sociedades de língua de sinais) que refletem as suas culturas contemporâneas;

5. Pesquisadores devem garantir que as opiniões e percepções do grupo de referência fundamental (o grupo de língua de sinais) apareçam em todo o processo de validação e avaliação dos dados na medida em que os termos de referência das comunidades de língua de sinais forem levados em conta;
6. Pesquisadores devem negociar dentro e entre os grupos de língua de sinais para estabelecer processos adequados de análise e determinar os critérios para decidir como atender os imperativos culturais, necessidades sociais e prioridades da comunidade.

Os dois trabalhos anteriormente citam o trabalho de Dwyer (2006) que também, juntamente com de outros pesquisadores da área, serviu como base para a criação das diretrizes éticas para pesquisa linguística da Sociedade de Linguística da América (SLA)⁹.

Tendo como base várias diretrizes éticas, incluindo as diretrizes da SLA, a Sociedade de Linguística da Língua de Sinais (SLLS) lançou em 2013 sua declaração de ética¹⁰, conforme informado no site da SLA. O informativo do site menciona ainda que a declaração de ética da SLLS tem vários pontos importantes. Um deles é a ênfase no comportamento ético no contexto do campo pesquisado, o que atende a muitas reivindicações apontadas pelas comunidades, inclusive as comunidades surdas. O site cita o trabalho desenvolvido por Crippen e Robinson (2013) que apresenta a problemática das pesquisas éticas que não levam em consideração as necessidades das comunidades pesquisadas e aprofunda a discussão sobre a colaboração em pesquisa linguística documental. Outra questão importante levantada na declaração é o estresse sobre os problemas éticos que resultam de relações de poderes diferentes entre os participantes da pesquisa. Embora essas desigualdades resultem de situações históricas, elas podem continuar a produzir efeitos sobre relações de trabalho contemporâneas e protocolos de pesquisa. Menciona-se, ainda, que a Sociedade tem a intenção de criar uma comissão de ética para acompanhar as questões decorrentes da declaração. Por entendermos ser um documento fundamental para as discussões de ética em pesquisa linguística com línguas de sinais, ser um documento publicado

⁹ Disponível em http://www.linguisticsociety.org/files/Ethics_Statement.pdf

¹⁰ <http://www.slls.eu/index2.php5>

recentemente e estar disponível somente em inglês, o documento foi traduzido e incluído no ¹apêndice da pesquisa como forma de fornecer material para os pesquisadores surdos que tem o inglês como terceira língua.

Em uma pesquisa de formação de corpus a partir da rede, que chamaremos aqui de corpus-web, todos os princípios éticos devem ser reajustados para essa realidade, visto que muitos dos parâmetros pensados aqui se aplicam para a grande maioria de pesquisas linguísticas em que a criação, filmagem e elicitación dos dados também fazem parte do projeto. No corpus-web os dados estão prontos, à disposição na rede. Inicialmente tem-se um perfil de informantes, mas não se sabe exatamente quais informantes serão encontrados ali.

Como o corpus formado nesta pesquisa trata-se apenas de um corpus-piloto com o objetivo primeiro de fornecer dados sobre circunstâncias e fenômenos metodológicos envolvidos durante a formação de um corpus-web, algumas das demandas para a formação do corpus não se aplicarão aqui. Dessa forma teremos três abordagens de adequações às diretrizes éticas: aquelas que são direcionadas para a grande maioria das pesquisas em corpus linguístico em que há a geração dos dados da pesquisa; aquelas que se aplicaram ao corpus-piloto da pesquisa e aquelas propostas pelos resultados da pesquisa e que podem servir de orientação para a formação de corpus-web.

Tratemos um pouco mais sobre a questão da relação estabelecida entre os participantes da pesquisa. Segundo Dwyer (Ibid), essa questão deve estar presente já na fase de planejamento. É preciso nessa fase definir como se dará a relação ética e legal entre todos os participantes do processo de documentação, pois essas relações irão desempenhar uma importante função ao determinar a direção e o escopo da pesquisa. Em geral há interesses subjacentes às instituições ou agências de fomento da pesquisa. O pesquisador deve pensar aqui também nos interesses das comunidades envolvidas na pesquisa.

A definição do Status do pesquisador nesse momento é importante para estabelecer o planejamento de sua pesquisa. Há questões diferentes implicadas se o pesquisador está trabalhando na condição de estar fora da comunidade (outsider) ou de pertencer a ela (insider):

The researcher might be an insider (i.e. accepted as a member of that community) or an outsider (from a distant community, whether in that country or in another). These roles are gradient rather than absolute, since a foreign researcher and a native speaker from a distant community may both be considered “outsiders” from the community under investigation. A local researcher often assumes multiple insider/outsider roles: it is often the case that a researcher is part of the ethnolinguistic group, but not or no longer from the particular community. In this situation, that researcher is both an insider and an outsider. The distinction may be relevant for research planning, as it often facilitates research to work with a person from the actual community under investigation. (DWYER, 2006, p.36)

Enquanto outsider, o pesquisador deve então nesse momento planejar a relação que será estabelecida com a comunidade. A possibilidade é de uma relação de parceria que beneficie tanto o pesquisador quanto a comunidade. Conforme Harris, Holmes & Mertens, (2009, p. 111) apud Hochgesang et al 2010:

In general, we share the opinion that ... “the formation of partnerships with researchers and the Sign Language communities is an important step in addressing methodological questions in research” (Harris, Holmes & Mertens, 2009, pp. 111)

O mapeamento desse contexto de relações entre informantes, pesquisadores e instituições que é apresentado por Dwyer pode ser visto de forma mais clara no quadro adaptado pela mesma:

Figura 2 – Quadro de Dwyer dos participantes da pesquisa linguística e suas relações

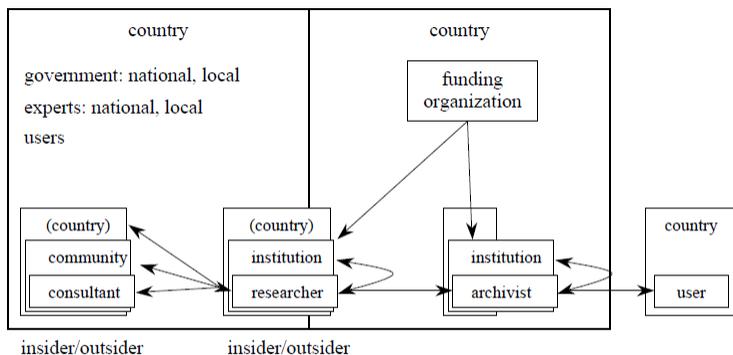


Figure 1. Participants in linguistic fieldwork (adapted from Hiß 2001, Wittenburg 2001–2004)

Fonte: DWYER, A.M. (2006). **Ethics and Practicalities of Cooperative Fieldwork and Analysis**. In J. Gippert, N.P. Himmelmann, & U. Mosel (Eds.), *Essentials of Language Documentation*. Berlin: Mouton de Gruyter, pp. 31—66

Para uma definição clara das partes envolvidas é necessário explicitar o uso que será feito dos dados. Isso porque a plataforma caracteriza a permissão legal baseada no uso. O uso para fins educativos, crítica, notícias, ensino, bolsa de estudo ou pesquisa, pode ser considerado aceitável. Usamos o termo *pode*, porque uma reivindicação desse tipo é definida por uma interpretação jurídica, que depende de alguns fatores. Há quatro fatores principais que são indicados como fatores de uso aceitável, conforme indicado por seção específica na plataforma¹¹ e discutidos aqui:

1. A finalidade e o caráter do uso, incluindo se tal uso é de natureza comercial ou se presta a fins educativos que não visem à obtenção de lucro

Os tribunais geralmente concentram-se no fato de o uso ser "transformativo" ou não. Em outras palavras, se o uso acrescenta uma nova expressão ou um novo significado ao original ou se meramente copia o original.

¹¹ <http://www.youtube.com/yt/copyright/pt-BR/fair-use.html>

2. A natureza da obra protegida por direitos autorais

Uso do material de obras primordialmente factuais é mais provável de ser aceitável do que o uso de obras puramente ficcionais.

3. A quantidade e a substancialidade de partes usadas em relação à obra protegida por direitos autorais como um todo

Usar parcelas pequenas do material de uma obra original é mais provável de ser considerado como uso aceitável do que usar parcelas grandes. No entanto, até mesmo uma parcela pequena pode ter um peso decisivo em algumas situações de uso aceitável, caso a parte em questão constitua o "cerne" da obra.

4. O efeito do uso sobre o mercado em potencial ou sobre o valor da obra protegida por direitos autorais. Usos que prejudiquem o lucro do proprietário dos direitos autorais sobre sua obra original são menos prováveis de serem usos aceitáveis. Tribunais já abriram exceções para esse caso quando envolve paródia.

A definição do caráter do vídeo é importante aqui para determinar se sua reprodução se enquadra em todos os quatro itens mencionados. Veja que o uso aceitável referido é para obras criativas e fixadas em um meio tangível. Ideias, fatos e processos não estão sujeitos a direitos autorais, conforme destaca o site¹². Se um usuário publica um fato corriqueiro do seu dia a dia, um depoimento ou algo do tipo cotidiano, não há direitos autorais envolvidos. Se terceiros aparecem no vídeo, ou se os vídeos são postados sem a autorização daqueles que aparecem ali, o que pode ser reivindicado por esses é o direito à privacidade. Exemplos desse tipo são citados como reivindicações jurídicas possíveis no youtube por Santos (2008), ao discutir o direito autoral na era digital. Cita a autora um primeiro caso, tratando de indivíduos que foram filmados sem consentimento em momento íntimo e o vídeo depois publicado no youtube, o que levou inclusive ao bloqueio do acesso ao canal no Brasil por um dia, após determinação judicial. O segundo relacionado a um filme,

¹² <http://www.youtube.com/yt/copyright/pt-BR/what-is-copyright.html>

produto criativo com direitos autorais, publicado na plataforma. Da mesma forma os vídeos cotidianos não são trazidos para a discussão por Serrano e Paiva (2008) que, ao encontro das reflexões propostas por Santos (Ibid), centralizam as discussões aos materiais postados sem autorização e aos detentores de direitos autorais. Esses dois tipos de materiais também são citados como passíveis de reclamação jurídica pela lei ⁹14.853/2013 ao dispor sobre a gestão coletiva de direitos autorais e classificar as obras que nela se enquadram.

Por natureza, então, podemos isentar as produções cotidianas, de autoria própria e sem a exibição de imagens de terceiros não autorizadas, de direitos autorais. No entanto a isenção legal não implica isenção ética. Se os vídeos são indicados no canal do pesquisador através de playlists – listas de vídeos organizados que remetem ao endereço da postagem original, não há necessidade de solicitação de autorização, pois não há violação de direitos autorais. Ao publicar o vídeo o usuário tem a opção de classificá-lo de três formas: vídeo privado - apenas para sua visualização e daqueles os quais o usuário informar email; não-listado - para visualização apenas das pessoas que tiverem o link informado pelo usuário e que não aparecem como resultado nas listas de buscas e; público - para visualização de qualquer pessoa com acesso a internet, podendo ser listado em playlist's de outros canais. Essa possibilidade de uso é útil quando o pesquisador necessita do corpus temporariamente, como para ministrar uma aula ou um curso, ou para uma pesquisa individual. O risco dessa utilização é que o controle dos dados não é feito pelo pesquisador. O proprietário do vídeo pode retirá-lo ou bloquear o acesso ao seu conteúdo a qualquer momento. Se os vídeos serão reproduzidos em outro material concreto que não aquele vislumbrado pelo seu autor para publicação (congressos ou artigos, por exemplo), embora possam ser considerados legais por uso aceitável, não são considerados uso ético devido a alta exposição de imagem que a língua de sinais proporciona e entram em confronto às reivindicações das comunidades surdas atualmente. Concluímos que qualquer utilização que não seja a reprodução em sala de aula e a pesquisa doméstica deve solicitar autorização de uso de imagem. Esse também parece ser o entendimento de Reppen (2005) que menciona:

In all cases, before collecting texts, it is important to have permission to collect them. When collecting texts from people or institutions, it is essential to get consent from the parties involved. The rules

that apply vary by country, institution and setting, so be sure to check before beginning collection. There are texts that are considered public domain. These texts are available for research and permission is not needed. Public domain texts are also available for free, as opposed to copyrighted material, which in addition to requiring permission prior to use may also have fees associated with it. Even when using texts for private research, it is important to respect copyright laws. This includes material that is available online. (REPPEN, 32)

Ao acessar 57 mil textos na internet para a compilação de um corpus de 45 milhões de palavras da língua portuguesa, Davies e Ferreira (2006) informam que por motivos de direitos autorais não podem disponibilizar os textos na página do corpus e que eles só podem ser acessados por seus endereços diretamente na internet. Os autores disponibilizam no site uma lista com o nome dos autores dos textos e o local onde os textos podem ser encontrados.

Figura 3 – Lista com as fontes do Corpus do Português

A	B	C	D	E	F	G
1	2	3	4	5	6	7
textID	genre	#palav	titulo	autor	data	fonte
26214	Bras-Fic	41742	Beltrão, Luiz	A Greve dos Desempregados	1984	SCANNED
26224	Bras-Fic	39392	Comparato, Doc	A Guerra das Imaginações	1997	SCANNED
26216	Bras-Fic	40579	Callado, António	A Madona de Cedro	1957	SCANNED
26260	Bras-Fic	40455	Vieira, José Geraldo	A Mais que Branca	1974	SCANNED
26248	Bras-Fic	42339	Penna, Cornélio	A Menina Morta	1958	SCANNED
26249	Bras-Fic	42091	Queiroz, Dinah Silveira de	A Muralha	1964	SCANNED
26239	Bras-Fic	39704	Montello, Josué	A Noite sobre Alcântara	1978	SCANNED
26226	Bras-Fic	42225	Costa, Eduardo Alves da	A Sala do Jogo	1989	SCANNED
26052	Port-Fic	2243	A CONSOADA	ABEL BOTEELHO		SCANNED
26008	Port-Fic	2526	Trânsito Impedido	AFONSO BOTEELHO		SCANNED
26003	Port-Fic	2471	Hora de Folga	AFONSO RIBEIRO		SCANNED
26122	Port-Fic	57165	O homem sem nome	Aguiar, João	1986	SCANNED
26117	Port-Fic	43589	Os incuráveis	Agustina Bessa Luis	1982	SCANNED
26015	Port-Fic	2405	O Novo	AGUSTINA BESSA LUIS		SCANNED
26120	Port-Fic	43225	Margem Norte	Alexandre Cabral	1979	SCANNED
26059	Port-Fic	1826	O CAGADO	ALMADA-NEGREIROS		SCANNED
26026	Port-Fic	2443	PONTA TENETE	ÁLVARO GUERRA		SCANNED
26284	Bras-Fic	42688	Memorial de um Passageiro de B	Amadeu Amaral	1921	http://alecncm.inf.ufsc.br/bdnuapl/
26285	Bras-Fic	2749	Novela e Conto	Amadeu Amaral		http://alecncm.inf.ufsc.br/bdnuapl/
26137	Port-Fic	38529	A Máscara e o Destino	Amorim, Antonio Guedes de	1944	SCANNED
25682	Bras-Acad	28208	O ensino do conceito de tempo	André Ferrer Pinto Martins		http://www.nilc.icmc.usp.br/facioweb/corpora.htm
25669	Bras-Acad	15906	Análise de redes metabólicas em	Andreas Karoly Gombert		http://www.nilc.icmc.usp.br/facioweb/corpora.htm
25684	Bras-Acad	18515	Uso de Redes Neurais para Corre	Andréia Gentil Bonfante		http://www.nilc.icmc.usp.br/facioweb/corpora.htm
26119	Port-Fic	40757	Os nos e os Laços	Antonio Alçada Baptista	1985	SCANNED
26077	Port-Fic	2462	Os Dois Barbeiros	ANTÔNIO ARNAUT		SCANNED
26278	Bras-Fic	12484	Cortos Avulsos	Antônio Castilho de Alcântara Machado	D'Oliveira	http://alecncm.inf.ufsc.br/bdnuapl/
26280	Bras-Fic	16398	Brás, Baviça e Barra Funda	Antônio Castilho de Alcântara Machac	1927	http://alecncm.inf.ufsc.br/bdnuapl/
26281	Bras-Fic	14871	Laranja-da-China	Antônio Castilho de Alcântara Machac	1928	http://alecncm.inf.ufsc.br/bdnuapl/
26282	Bras-Fic	17820	Mana Maria	Antônio Castilho de Alcântara Machac	1935	http://alecncm.inf.ufsc.br/bdnuapl/

Fonte: <http://www.corpusdoportugues.org/help/cdp.xls>

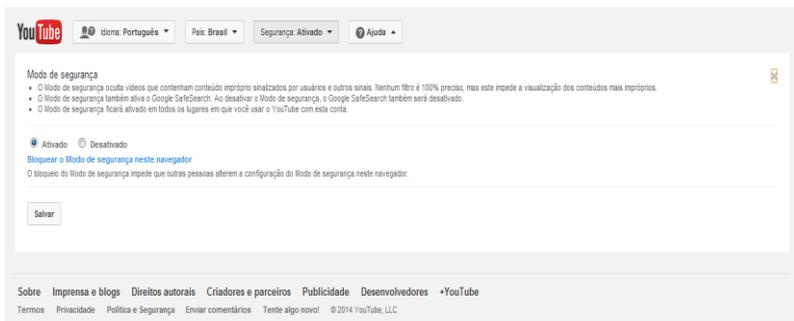
Eles caracterizam o trabalho como “uso aceitável” do corpus por apresentar apenas parte dos textos no resultado final:

Our corpora contain hundreds of millions of words of copyrighted material. The only way that their use is legal (under **US FAIR USE LAW**) is because of the limited "Keyword in Context" (KWIC) displays. It's kind of like the "snippet defense" used by Google. They retrieve and index billions of words of copyright material, but they only allow end users to access "snippets" of this data from their servers. (DAVIES E FERREIRA, 2006)

Essa poderia ser uma alternativa de publicação de obra com um número extensivo de fontes, o que inviabilizaria a pesquisa, que também foi utilizada por Schallemburger e tem sido usada por outros pesquisadores ao citar o caminho para o vídeo original na pesquisa. Há de se considerar, no entanto, a possibilidade de exclusão do vídeo da rede que retiraria da pesquisa científica a condição de verificabilidade.

Para a solicitação de autorização, o primeiro passo, então, é definir a autoria do material. O vídeo em questão pode ser resultado de uma re-postagem por alguém que se passa por autor do vídeo e ter sido sinalizado como violação de privacidade. Durante a pesquisa detectamos que a plataforma tem duas formas de suspender a exibição de um vídeo por violação dos direitos autorais ou privacidade; a primeira através de uma notificação de violação de direitos autorais feita pelo próprio autor do vídeo, ou alguém que o conheça e enviada pelo link disponível na página. Após essa notificação o vídeo é automaticamente suspenso. A segunda através de uma detecção automática feita pelo site denominada "content ID". Há também a possibilidade de excluir da busca vídeos considerados impróprios pelos usuários da plataforma. Isso acontece quando algum usuário percebe que há conteúdo particular utilizado no vídeo, conteúdo pornográfico, ofensivo, preconceituoso ou do gênero. Como forma de iniciar uma busca para o corpus por vídeos pré-selecionados que não tenham reivindicações de privacidade ou direitos autorais, é possível utilizar um recurso da plataforma denominado modo de segurança, conforme vemos na figura abaixo:

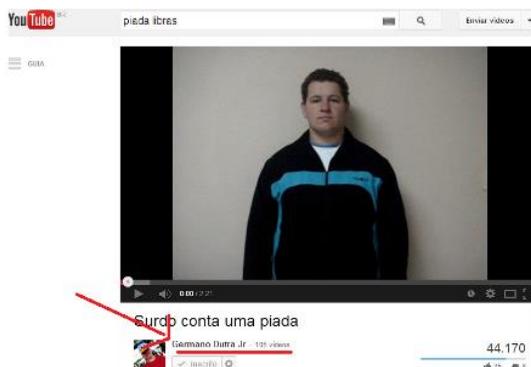
Figura 4 – Tela de Ativação do modo segurança na plataforma youtube



Fonte: parte inferior da plataforma www.youtube.com.br

Após ativado o modo de segurança, e o vídeo pré-selecionado por atender às necessidades da pesquisa, é necessário acessar o canal postador do vídeo e determinar sua autoria. A figura abaixo mostra como, após a seleção de um vídeo, pode-se acessar o canal do postador do mesmo:

Figura 5 – Localizando o postador de um vídeo na plataforma youtube



Fonte: <https://www.youtube.com/watch?v=697SwM6U-9I>

Clicando no link do nome conseguimos visualizar a página principal do mesmo e, em seguida, clicando em “sobre” obtemos mais informações do canal. Há um link disponível para que possamos enviar

mensagens para o proprietário do canal. É possível que na página inicial o proprietário também inclua outras formas de contato ou outras páginas de propriedade do mesmo indivíduo, como no exemplo a seguir:

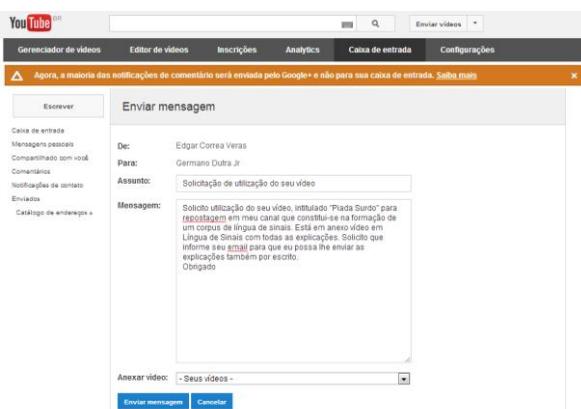
Figura 6 – Obtendo mais informações sobre o autor de um canal na plataforma youtube



Fonte: https://www.youtube.com/channel/ucyuueixfcg0rxp8w_hqs2fq

Após clicar em “sobre”, a tela de informações do canal, com o link de envio de mensagem de contato com o proprietário é exibida. Note que é possível enviar um vídeo em anexo à mensagem, o que permite que todos os contatos sejam feitos em libras.

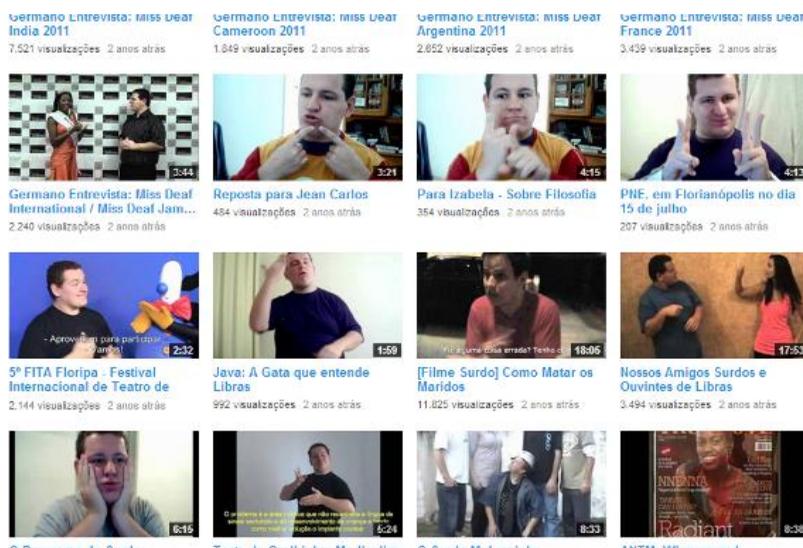
Figura 7 – Entrando em contato com o postador do vídeo



Fonte: <https://www.youtube.com/user/germaju/about>

Ao visualizar os vídeos postados pode-se atestar a autoria do canal. Analisando um dos canais que hospeda vídeos selecionados na pesquisa, nota-se pela quantidade de vídeos: 105, tempo de atividade do canal: desde 2007, quantidade de visualizações: mais de meio milhão - 595.514 e tipos de vídeos enviados: quase todos com exibição da mesma pessoa, que se trata de um canal de autoria original.

Figura 8 – Visualização de vídeos postados em um canal específico da plataforma youtube.com



Fonte: <https://www.youtube.com/user/germaju/videos>

Canais que postam vídeos de autoria de outras pessoas, em geral, não têm tantas visualizações e não são mantidos por muito tempo. Esses canais são “derrubados” pela plataforma após várias notificações de uso indevido. Mesmo assim, os vídeos selecionados podem ser monitorados antes da re-postagem. Um vídeo que infringe os direitos autorais em geral apresenta um curto tempo de exposição. Monitoração realizada por Serrano (2009, p.10) apresentou um tempo de 16 dias para a derrubada de um vídeo nessas condições.

Após a autorização e seleção dos vídeos, seguindo o primeiro princípio ético estabelecido por Dwyer e sistematizado também no primeiro item da declaração de ética da SLLS quanto à responsabilidade

para com os informantes, verificamos se as informações dadas pelos informantes não poderiam comprometê-lo perante a comunidade, isso é, se o autor não menciona terceiros em seu vídeo, ou informações consideradas segredos culturais ou que ofendam a comunidade.

Durante a pesquisa, procedimentos semelhantes foram encontrados na formação de outros corpora de língua de sinais. No *Corpus NGT - Sign Language of the Netherlands* – língua de sinais dos Países Baixos (Crasborn, Zwitserlood, and Ros, 2008¹³) a possibilidade de “mascarar” os vídeos logo foi descartada por não permitir que elementos importantes da gramática da língua de sinais que aparecem no rosto do sinalizador fossem observados, conforme ilustra Crasborn (2010):

Figura 9 – Tentativa de mascarar o rosto do sinalizador no Corpus NGT



Fonte: CRASBORN, O. **What Does “Informed Consent” Mean in the Internet Age? Publishing Sign Language Corpora as Open Content.**

Sign Language Studies, 10-2, 2010, p. 283. Disponível em

<http://www.ru.nl/publish/pages/515325/crasborn2010_sls_informedconsent.pdf>

Uma alternativa seria a conversão da imagem em desenhos de 3 dimensões, denominados de avatar. Tecnologias para essa finalidade têm sido desenvolvidas, conforme cita o autor mencionando o projeto Dict-a-Sign¹⁴. Como a operacionalidade do projeto ainda não está disponível em software aberto e seus recursos ainda precisam ser aperfeiçoados, não é possível utilizá-lo de forma ampla e funcional atual-

¹³ Disponível em <http://www.ru.nl/corpusngtuk/>

¹⁴ Disponível em <http://www.dictasign.eu/>

mente. Logo, as estratégias de privacidade incluíram orientações quanto ao conteúdo e menção de terceiros e explicitação clara da publicação dos dados conforme explica o autor:

Several measures were taken to ensure that the signers were aware of the nature of the publication. In the selection of participants for the recordings, it was explicitly mentioned that the goal of the project was to publish video recordings online for public access. When people arrived at the recording site, the Deaf assistant who was in charge of the recording session reminded the participants of the fact that all of the recordings would in principle be made available online. In addition, everyone was asked to be careful in what they talked about and to avoid gossip and, where possible, the mention of any names. Although language elicitation tasks were already careful to avoid personal stories, there was quite some leeway in exactly how discussions took place that might lead to privacy-sensitive remarks. ¹⁴(CRASBORN, 2010, p. 283)

Observamos que grande parte do cuidado se deve à conscientização do meio de publicação dos vídeos, a rede. No caso de um corpus-web em que os vídeos já estão disponibilizados, o informante deve ser orientado que a republicação do vídeo em um canal de pesquisa pode atrair uma faixa de público específico para o vídeo: a de pesquisadores e acadêmicos, que talvez não seria a mesma na publicação original do vídeo disponível a um público aleatório. Embora o mesmo público já tivesse acesso aos vídeos em seu canal original, a organização em um canal de pesquisa, bem como a disponibilização de metadados, arquivos de transcrição e frames do vídeo pode estimular o uso em pesquisa científica e os autores devem estar conscientes disso. Outro aspecto a ser observado é quanto à possibilidade de exclusão dos vídeos. Em seu próprio canal o proprietário gerencia o acesso, disponibilização de visualização e retirada do vídeo em qualquer tempo. Repostado, o gerenciamento do vídeo passa para as mãos do pesquisador. Em outras palavras, será preciso contato com o pesquisador caso o proprietário do vídeo deseje suspender sua exibição em algum tempo.

No Corpus NGT, mesmo após as orientações e gravação dos vídeos, os informantes receberam cópias de suas gravações antes da publicação para que pudessem assistir e avaliar novamente questões de privacidade no conteúdo de sua sinalização. Os pesquisadores retiraram todos os trechos iniciais das filmagens em que os informantes se apresentavam e falavam de si. Mesmo assim, após a reanálise, cinquenta sessões foram excluídas a pedido dos sinalizantes.

Assistindo aos vídeos notamos que em alguns, os autores se referiam a terceiros ou mesmo o vídeo era direcionado a alguém específico. Quando a reincidência era muito alta o vídeo foi retirado da amostra. Aqueles que eram considerados muito importantes por outras características que o tornavam protótipo do gênero (qualidade do vídeo, clareza da filmagem, função clara do vídeo), foram editados de forma a se retirar o sinal do terceiro informado e também o nome do terceiro informado nos dados do vídeo. Foi o que aconteceu no vídeo a seguir em que se retiraram as informações de terceiros.

Figura 10 – Retirando informações nominais de vídeo postado no youtube



Fonte: <https://www.youtube.com/watch?v=8VifwuwE3CQ>

Alguns canais, como o canal acima, apresentam o nome verdadeiro do autor, em outros apenas o “nickname”, um apelido utilizado na internet, aparecia como identificação dos sinalizantes. Para todos no

entanto foi solicitado no termo de consentimento a utilização do nome verdadeiro no corpus, além é claro das informações de metadados.

Quanto a elaboração do termo de consentimento Crasborn (2010) levanta algumas questões importantes:

The extent to which subjects are properly informed may not be as easy to judge as one would think (Schultz, Pardee, and Ensinck 1975). In the case of sign language corpora, one should ask first of all to what extent deaf people with varying levels of literacy in the researcher's spoken language (or in any spoken language for that matter) are aware of the status and impact of a consent form. Although the statements on the consent form may be explained to them in sign language (as was the case for the Corpus NGT), this does not count as a voluntary and well-informed decision if people are not fully literate and cannot comprehend the impact of a short, written document like a consent form (pág.285)

Algumas estratégias são propostas aqui. Na condição de insider na comunidade, o pesquisador poderá selecionar na amostragem informantes bilíngues ou que tenham familiaridade com a pesquisa científica (professores e informantes com nível superior, por exemplo). Na condição de outsider poderá trabalhar em parceria com um pesquisador que pertença à comunidade. No entanto, conforme a citação anterior, mesmo o termo estando em língua de sinais, isso não significa que o informante compreenda o impacto da utilização do vídeo na internet. Seria necessário um nível básico de letramento digital, conforme pontua ainda na sequência o autor:

First of all, knowledge of the Internet and its developments is needed. If one has no experience browsing web pages from all over the world (but only those from one's own country), it is hard to decide whether or not video recordings of oneself should be made available to the whole world. Alt-

though this was not explicitly checked, it was the impression of our Deaf assistant that even the older signers in the Corpus NGT were acquainted with the Internet and would be able to comprehend the meaning of “worldwide availability.”

Aqui há uma particularidade do corpus-web: em geral os vídeos são postados por seus proprietários, o que exige um conhecimento mínimo de manuseio da internet e pressupõe um melhor entendimento do autor quanto ao impacto das informações veiculadas na rede. Claro que a interface fácil da plataforma e aplicativos simplificados de smartphones que enviam vídeos gravados automaticamente podem fazer com que usuários que não estejam tão bem familiarizados com a internet insiram vídeos no canal. Isso sem falar também que um terceiro pode inserir um vídeo de autoria de outra pessoa como é comum com vídeos de surdos idosos em eventos, bate-papo ou entrevistas informais. Para isso, como estratégia, o pesquisador deve selecionar na amostragem, vídeos de canais que sejam de autoria própria do postador, que possuam um número significativo de vídeos publicados e um tempo razoável de operação na plataforma. Não sendo possível, deve buscar explicar cuidadosamente aos protagonistas dos vídeos os possíveis desdobramentos da publicação de forma clara e que possibilite a compreensão da realidade virtual.

Implicações ainda podem ser pensadas quanto a manter os dados na plataforma por tempo indeterminado. Não é possível prever as implicações tecnológicas da disponibilização de imagens na rede, como a possibilidade de reconhecimento automático de identidade a partir de imagens por softwares cada vez mais modernos. Um bom exemplo é apresentado por Crasborn (2010) quando se refere ao fato comum de olharmos para fotos de 20 anos atrás e nos envergonharmos de nossa aparência, estilo e aspectos apresentados ali. O mesmo poderia ocorrer com os vídeos. A melhor alternativa aqui é possibilitar ao informante, a qualquer momento, solicitar a retirada de seus vídeos do canal na plataforma.

Os principais itens mencionados no corpus NGT foram:

1. Aceito ser gravado em fita de vídeo para o projeto Corpus NGT.

2. Concordo que as gravações de vídeo sejam disponibilizadas através da Internet e dessa forma sejam usadas para pesquisa, ensino etc. As gravações serão livremente disponibilizadas e podem ser usados sem custo. Os vídeos não podem ser utilizados para fins comerciais ou qualquer forma de obtenção de dinheiro. "Trabalhos derivados" podem ser feitas como legendagem dos filmes ou usar imagens das gravações para páginas da rede ou apresentações. Estas obras derivadas também não podem ser utilizadas para qualquer fim comercial que vise a obtenção de lucro.
3. Concordo que estes dados sejam disponibilizados por tempo indeterminado (na Internet ou através de outros meios de comunicação).
4. Eu concordo em ser agradecidas por nome nas publicações e no site do projeto.

Considerando todos os aspectos abordados e as adaptações necessárias para a pesquisa elaborou-se o termo de consentimento em duas versões, a versão em língua portuguesa e em Libras. O termo foi anexado à pesquisa.

Ainda tratando dentro do escopo das relações estabelecidas na documentação linguística, nas diretrizes éticas de Dwyer (2006), item 3, vimos a orientação para "buscar fazer o bem a comunidade" além do bem à ciência. Essa posição foi reforçada na sistematização de diretrizes de Harris, Holmes & Mertens (2009) apud Hochgesang et al (2010) que mencionam no item 6, a necessidade de decidir junto com a comunidade quais são as demandas urgentes que melhor podem ser atendidas na pesquisa. A declaração ética da SLLS oficializou essas orientações em um sub-item específico, o de número 2: "responsabilidade para com as comunidades de surdos".

Dudis, Mathur e Mirus (2010), na mesma linha dos autores citados anteriormente, trouxeram algumas propostas sobre a elaboração de um corpus adequado às perspectivas da comunidade surda. Elaborando um planejamento piloto no encontro de jovens surdos nos Estados Unidos, os pesquisadores entrevistaram um grupo de cinquenta jovens

em que 82% aprenderam ASL antes dos cinco anos de idade. Os entrevistados responderam perguntas abertas quanto ao envolvimento da comunidade surda e aos projetos de criação de corpus.

Algumas perguntas também questionavam sobre a importância e o estado atual da ASL na visão dos entrevistados:

- Você acha que haverá mais pessoas usando ASL em 300 anos?
- Você acha que a inclusão de crianças surdas mais afetará o número de pessoas que utilizam ASL?
- Você acha que a tecnologia médica atual já está reduzindo o número de pessoas surdas que utilizam ASL?
- A gramática da ASL deve ser ensinada aos alunos surdos na escola em paralelo ao ensino de Inglês? Deveria haver uma campanha para preservação da ASL da mesma forma como foi feita pela associação nacional de surdos no início dos anos 1900?

Os autores concluem com reflexões acerca da necessidade de um envolvimento mais efetivo da comunidade surda nos projetos de formação de corpus e a necessidade de um retorno mais significativo para essas comunidades.

Baseado no trabalho proposto, elaboramos um questionário adaptado a partir do questionário de Dudis, Mathur e Mirus (2009). O questionário tinha o objetivo de analisar as necessidades da comunidade surda brasileira em relação à produção de corpus de línguas de sinais. Para isso foram selecionados 20 acadêmicos surdos. Priorizaram-se indivíduos surdos que possuíam algum tipo de atuação voluntária em organizações civis de apoio a comunidade surda, como associações de surdos, escritórios da federação nacional de educação e integração de surdos (Feneis), escolas e centros de apoio à educação de surdos. O questionário foi dividido em cinco partes: sondagem individual; apresentação de corpus de línguas de sinais; questões sobre o uso de corpus da língua; questões sobre utilidade da pesquisa linguística para a comunidade surda e questões sobre corpus e gêneros textuais. O questionário pode ser visualizado no anexo 2 da pesquisa.

Pensando ainda em uma forma de retorno à comunidade, estabelecemos parceria com a Federação Nacional de Educação e Integração

de Surdos para auxiliar na tradução e interpretação de demandas consideradas urgentes para a comunidade no projeto político por uma educação bilíngue para surdos. Foi inserida uma pasta no corpus com traduções e materiais produzidos em língua de sinais pelo pesquisador a partir da demanda apresentada pela federação, além de um espaço para a publicação da versão da dissertação em libras que embora não seja obrigatória pela Universidade, foi produzida como forma de devolução à comunidade surda dos resultados de pesquisa. O questionário de sondagem junto à comunidade também foi postado nesse espaço. Os espaços e materiais produzidos podem ser visualizados a seguir:

Figura 11 – Materiais produzidos para em auxílio a federação nacional de educação e integração de surdos



Fonte:

https://www.youtube.com/channel/ucqdw2sal29t_avjhzxzaoa/videos

3 QUESTÕES METODOLÓGICAS PARA A FORMAÇÃO DO CORPUS

Neste capítulo continuaremos tratando dos procedimentos adotados na formação do corpus-web. Dentro do escopo de pesquisas para formação de corpus audiovisual, faremos um breve apanhado dos procedimentos adotados para a pesquisa com vídeos na web, mostrando como foi feita a escolha da plataforma de pesquisa. Em seguida, discutiremos os procedimentos de seleção dos vídeos e escolhas dos gêneros textuais para compilação.

3.1 Da extração de vídeos na rede para a formação de corpus de línguas de sinais

Para a manipulação dos textos na internet a primeira orientação de Berber Sardinha (2004, p.46) é localizar as páginas que servirão como fonte, extrair os textos e salvá-los no computador. Em se tratando de coleta de grandes quantidades de textos, coletas sistemáticas e longitudinais, é necessário, conforme o autor, ferramentas que façam download automaticamente e que salvem as informações no computador. O autor sugere ainda o uso de offline browser para a realização desse trabalho, conforme vemos a seguir no quadro retirado da obra referida anteriormente:

Figura 12 – Tela de programa para extração e download de informações de texto de páginas da web

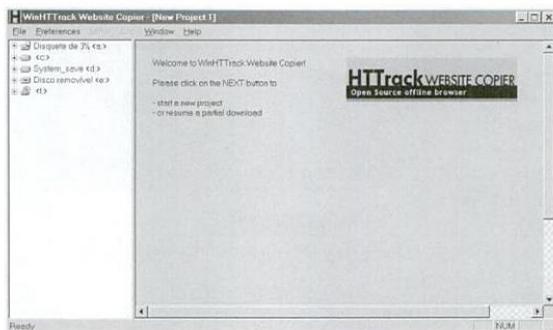


FIGURA 2.1

Tela inicial do *offline browser* Win HTTrack.

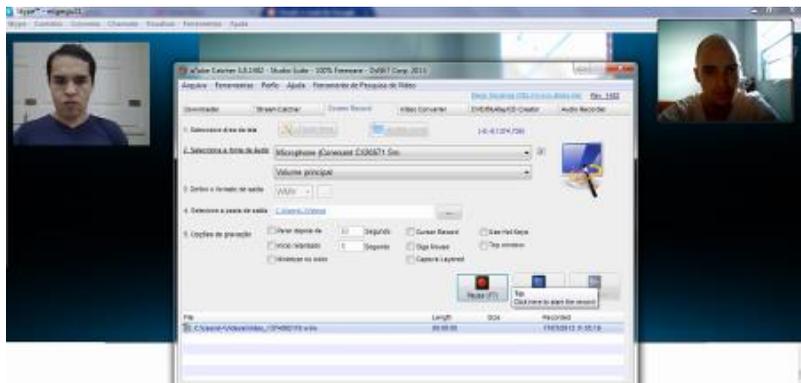
Fonte: BERBER SARDINHA, T. *Linguística de Corpus*. Barueri, SP: Manole, 2004, pág.46

Em geral, os vídeos postados no espaço virtual são gerenciados pelas páginas (web-sites) que os publicam ou pelo próprio postador, que utilizam de espaços de blogs, redes sociais e outras plataformas que dão a seus usuários essa possibilidade com pouco ou nenhum monitoramento. Como o pesquisador necessitará fazer a edição, visualização de frames e cópias de segurança e como os dados podem ser retirados da rede a qualquer momento, é necessário o uso de ferramentas de extração dos vídeos. Para realização dessa tarefa inicial na pesquisa, levantamos através do motor de busca google.com os softwares livres mais utilizados para este trabalho. Analisando as dez primeiras respostas de busca vimos ser possível classificar em dois tipos as ferramentas para isso(2) programas que podem ser instalados em seu computador e que permitem realizar essa extração. Visitando cinco principais sites mais recorrentes para essa finalidade nos resultados da busca, vimos que o uso dos sites não atende a uma das principais demandas para a extração de grande quantidade de dados na internet conforme Berber Sardinha (2004): a extração automática de dados. Dessa forma nos limitamos a pesquisar programas que realizassem essa função, retirando os sites que apareceram no resultado da busca. Como resultado desse refinamento de busca, trinta e oito principais programas foram encontrados em dez sites especializados. Para fazer a seleção do programa a ser utilizado precisávamos

mos definir as características importantes que consideraríamos como critério na escolha. Os aspectos considerados para a escolha do programa foram: freeware registrado e com licença de distribuição; funcionalidade dos recursos; capacidade de extração automática de vídeos; agregar download e conversão do vídeo em um único programa e possuir orientações quanto aos direitos autorais da utilização do programa. Após a seleção de cinco principais programas que atendiam a esses requisitos, definimos o uso do programa atube catcher¹⁵, sobretudo por além de atender todos os critérios definidos, possuir uma interface de fácil utilização. Além do download de vídeos, o programa oferece a possibilidade de fazer extrações de qualquer fonte de vídeo, utilizando a ferramenta de captura de tela (função Screen Record). Caso o pesquisador deseje coletar dados de um informante surdo, por exemplo, que não esteja familiarizado com a gravação e inserção de vídeos na plataforma youtube, e além disso tratar-se de um informante que reside um local distante do local de residência do pesquisador, poderá solicitar um contato utilizando programas de video-chat muito difundidos e utilizados entre a comunidade surda como Facebook messenger, ooVoo, Skype etc.

Utilizando essa função o pesquisador poderá selecionar o quadro da tela que deseja capturar, ou mesmo a tela toda, e prosseguir com a coleta enquanto faz uma entrevista em língua de sinais por exemplo.

Figura 13 - Utilizando a ferramenta Screen Record para gravar uma entrevista em língua de sinais no Skype



Fonte: Tela de Visualização de conversa no Skype

¹⁵ Disponível para download em <http://atube-catcher.dsnetwork.com/video/>

Há, no entanto, algumas dificuldades de extração de vídeos quando o vídeo está numa configuração não muito comum ou compatível com a maioria dos softwares de extração de vídeos, como é o caso de arquivos SWF¹⁶. O programa também permite a extração automática de vídeos em formato SWF através da função Stream Catcher.

Definidas as técnicas de gravação é necessário definir as fontes de extração de vídeos. A internet dispõe de várias plataformas de hospedagem de vídeo, páginas da rede ou softwares onde os usuários podem distribuir seus vídeos. Alguns anexam esse serviço a outros, como as redes sociais, em que é possível postar arquivos, vídeos, imagens etc. Essas redes podem possibilitar a opção de compartilhamento privado dos vídeos. Outros funcionam apenas como motores de busca desses vídeos sem a possibilidade de hospedar os vídeos como o Singingfish¹⁷ e o Mobento¹⁸.

O programa atube catcher oferece também a possibilidade de busca dentro os principais serviços de hospedagem de vídeos¹⁹. Para a pesquisa, delimitou-se a utilização da plataforma youtube pelos critérios detalhados a seguir. Como membro da comunidade surda, percebemos a popularidade dessa plataforma dentre as demais utilizadas e mencionadas nesta pesquisa. Além disso, a plataforma é apresentada por diversos sites de ranking da categoria como o melhor site de hospedagem e compartilhamento de vídeo²⁰. É a plataforma que possui maior quantidade de vídeos (+ de 460 milhões) e a maior quantidade de aces-

¹⁶ SWF (Shockwave Flash), tipo de arquivo que comprime as informações de vídeo permitindo que possam rodar na rede com melhor transmissão. Muitos sites que baseiam sua interface em língua de sinais utilizam arquivos SWF, principalmente aqueles em que a quantidade de vídeos é extensa, como os dicionários de sinais.

¹⁷ A plataforma foi adquirida pela Aol e está disponível em <http://on.aol.com/>

¹⁸ Disponível em <http://www.mobento.com/>

¹⁹ Uma lista completa de serviços pode ser vista em <http://www.ranker.com/list/free-video-sharing-websites/sirsinister>

²⁰ Como pode ser visto em <http://video-share-review.toptenreviews.com/>

sos por dia (1 bilhão e 200 milhões), incluindo o melhor sistema de busca e filtragem online²¹.

3.2 Organizando o corpus de vídeo

Definidos os mecanismos de busca e o site que servirá como fonte de dados é necessário definir a organização dos vídeos que serão baixados para a formação do corpus.

Após a gravação, o vídeo é automaticamente enviado para a pasta pré-definida do seu computador que pode ser selecionada de acordo com a organização das pastas do seu corpus.

Quanto à organização de pastas, Berber Sardinha (2004) pontua que não há uma regra geral para isso. Há diferentes tipos de organizações que vão depender principalmente da finalidade do corpus e da pesquisa. No entanto, algumas recomendações importantes são dadas pelo autor:

- Os textos (considerar vídeos para essa pesquisa) devem estar em uma pasta principal em que só existam os textos do corpus, por exemplo, c:\corpus
- Deve ser criada uma subpasta que indique a versão atual do corpus, por exemplo 00
- As subpastas criadas devem refletir seu conteúdo, ou seja, tenham nomes indicativos do tipo de texto, assunto, modalidade, registro.
- Depois é aconselhável copiar toda a estrutura montada e salvá-la com o nome *temp* que tem a função de ser uma cópia de trabalho do corpus onde são feitas as modificações nos arquivos, transportados para a pasta 00 quando estiverem prontos para o uso.

²¹ Uma comparação técnica detalhada das plataformas que comprovam os argumentos mencionados pode ser visualizada em http://en.wikipedia.org/wiki/Comparison_of_video_hosting_services

Durante a realização da pesquisa, cerca de 100 vídeos em seus formatos originais foram baixados para um computador pessoal. Após a escolha do software para transcrição dos vídeos, o ELAN²², tivemos uma grande dificuldade de utilização dos vídeos no software por conta de incompatibilidade de codificadores de vídeos. Foi necessária então uma conversão dos vídeos para a extensão MPEG²³, a mais compatível com o software. Após a conversão dos vídeos, notamos que um índice grande dos arquivos apresentou erro, cerca de 10%. Logo, optamos por manter uma versão do arquivo original baixado, caso novos erros possam aparecer futuramente e exijam novas conversões quando o arquivo já não estiver disponível online. Prevendo a futura necessidade de compartilhamento dos vídeos com outros pesquisadores, pareceu ser mais adequada a manutenção de versões dos vídeos em diretórios separados: um diretório com os vídeos no formato convertido, para compartilhamento, e outro com os vídeos originais.

Com 100 vídeos originais e 100 cópias desses vídeos em formato *mpeg* na memória do computador pessoal, percebeu-se aumento da lentidão no processamento de informações do computador (com memória ram de 2gb, a média atual em computadores de uso pessoal com processadores duo). Esse fato somado à possibilidade de perda dos dados em alguns momentos sugeriu novos direcionamentos para a organização de corpus de vídeos além daqueles orientados por Berber Sardinha. Também notamos a necessidade de acesso a algumas informações de sinalização de vídeos ainda não transcritos no ELAN. Esses acessos necessitavam ser feitos de forma rápida sem que seja preciso assistir todo o vídeo novamente. Uma visualização em duas dimensões parecia atender melhor essa necessidade. Somada a possibilidade de compartilhamento com não usuários do ELAN que também necessitariam fazer esse tipo de manuseio propomos a manutenção de arquivos de frames do vídeo. Sistematizamos, portanto, as seguintes orientações adicionais para corpus de vídeo:

²² Disponível em <http://tla.mpi.nl/tools/tla-tools/elan/>

²³ Moving Picture Experts Group (MPEG) ou Grupo de Especialistas em Imagens Animadas. Esse foi um projeto desenvolvido por especialistas na área no final da década de 1980. Esse grupo desenvolveu um sistema de compactação de arquivos de filme para armazenagem e reprodução em computadores, sem excesso de tamanho ou peso. No entanto para fazer a compactação/codificação e posterior decodificação desses vídeos é necessário ter os codificadores e decodificadores específicos pré-instalados no seu computador.

- Todos os arquivos de vídeos baixados devem ser convertidos para o formato de extensão de vídeo compatível com o software escolhido para a transcrição;
- O diretório com os arquivos convertidos deve estar no local onde o software de transcrição está instalado: no computador, já que o software “roda” fazendo um link com os vídeos fora dele. A manutenção dos vídeos em diretório externo mostrou instabilidade no uso;
- O diretório temporário dos vídeos deve ser mantido em uma unidade externa, só ligada ao diretório central durante o uso;
- Os arquivos de frames devem ser mantidos junto aos arquivos de vídeo originais que permitam acesso rápido das informações dos vídeos.

A proposta de organização de ⁵Berber Sardinha pode ser visualizada a seguir:

Figura 14 - Proposta de organização de Berber Sardinha para diretório de pastas na formação de corpus coletado na rede

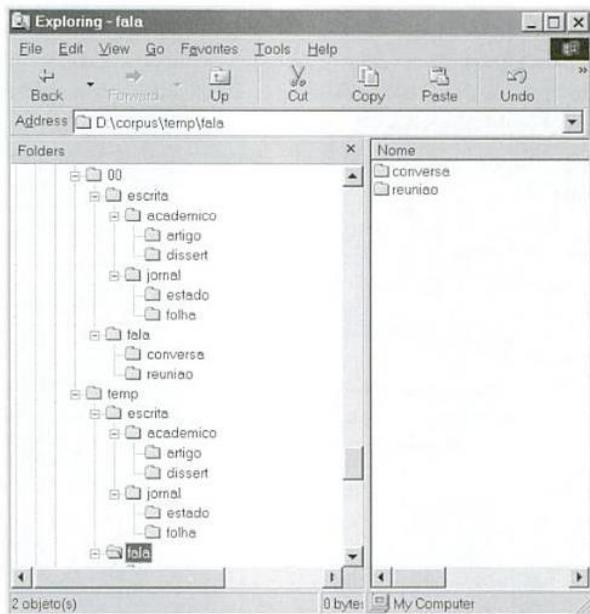


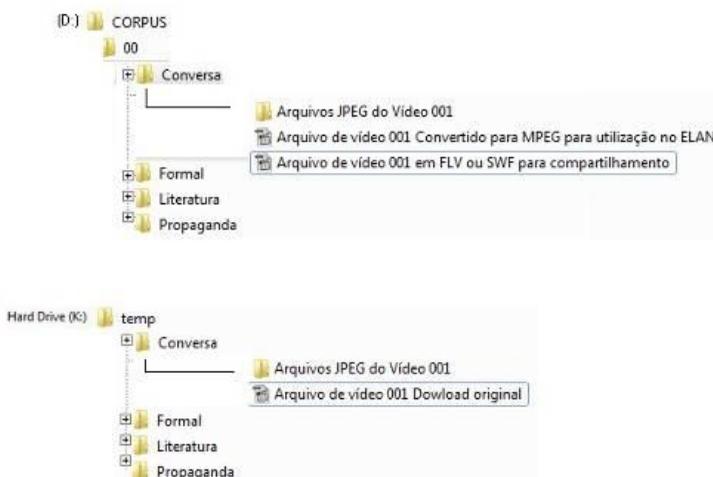
FIGURA 2.9

Proposta de estrutura de diretórios do corpus.

Fonte: BERBER SARDINHA, T. Linguística de Corpus. Barueri, SP: Manole, 2004, pág.73

Considerando as orientações e adequações propostas para um corpus de vídeo, uma organização poderia ser sugerida conforme a figura a seguir:

Fig – 15 - Proposta de organização de pastas para um corpus de vídeos



Fonte: Tela do diretório de computador pessoal do autor

Após a definição da sistematização da organização dos vídeos, tratemos dos procedimentos para tratamento individual de cada um dos vídeos do corpus.

3.3 Limpeza do corpus de vídeo

Os textos do corpus devem ser “limpos” após o seu armazenamento. Isso envolve, conforme Berber Sardinha, a formatação dos textos de acordo com o software que irá processá-los. Para a manipulação de textos escritos de um corpus no WordSmithTools por exemplo, um dos softwares mais utilizados para esse fim, recomenda-se que os arquivos estejam no formato .txt, o que significa que esses arquivos contêm apenas caracteres do teclado (letras, números e símbolos gráficos).

De fato, em se tratando de arquivo de vídeo e seus respectivos codec's²⁴ tem se apresentado como um elemento importante para a utilização de softwares processadores de vídeos em língua de sinais. É o que vemos em problemas apontados pelos usuários do software Elan, um dos principais softwares utilizados para transcrição de língua de sinais e publicados por Sloetjes, Wittenburg e Somasundaram (2011, p. 23):

ELAN is mainly written in the Java programming language, which should guarantee cross platform code and binaries, but for media play back it connects to one or more of the available “native” media frameworks. One of the advantages for the user of this approach is that many media file formats are supported (the down side is that problems with installed codecs are inherited as well). Where in the early years of ELAN only mpeg-1 files were “allowed” for video and wave files for audio, now any file type that is supported natively or by additionally installed codecs is accepted. In fact this has implications for interoperability as well: while the emphasis mostly is on the possibilities of exchanging the annotation data, often the converted data need to be seen in combination with the source of the data, the media files. Using “exotic” media file types potentially reduces the interchange success rate (and along these lines one could say that an application allowing the user to pick any type of file does the same). This is the common trade-off between convenience for the user and consequences for interoperability (pág. 1)

²⁴ Codificadores e decodificadores, são programas que codificam e decodificam arquivos de mídia, favorecendo compactação para armazenagem e descompactação para visualização. Uma das soluções encontrada pelos linguistas de corpus que utilizam computadores pessoais para processamento de corpora é a instalação de pacotes de codec's que possuem quase todos os codificadores/decodificadores da maioria dos formatos de vídeo encontrados no mercado. Pacotes do tipo podem ser encontrados em http://codecguide.com/download_mega.htm

Os problemas resumidos por Sloetjes et al, em relação a Codec's podem ser exemplificados por aqueles apontados por Bickford (2005) quanto à necessidade de utilização de arquivos no formato .mpeg com as especificações de Codec's e por Leite (2008) que apesar de utilizar vídeos na extensão .mpeg e .mov teve problemas com operacionalidade de vídeos no programa por conta da não utilização de codec's específicos.

Ainda tratando da limpeza dos textos, Berber Sardinha (2004) orienta sobre a necessidade de usos de softwares para uma limpeza eficiente dos textos no sistema operacional Windows, que retire caracteres desnecessários, transforme acentos em caracteres simples, efetue operações como encontrar, substituir e formatar e todas as demais tarefas de pré-processamento de textos. Em um corpus de vídeos idealizado e configurado pelo pesquisador, questões de imagem, angulação e questões que irão refletir diretamente na qualidade da manipulação dos dados podem ser previamente pensadas.

Já de início convém destacar a dificuldade de aplicação de corpora de vídeos da rede para pesquisas específicas que investiguem a produção de sinais em co-ocorrência com os articuladores não-manuais ou que dêem ênfase a aspectos fonéticos e morfológicos. Mesmo em pesquisas eliciadas e experimentais frequentemente a produção necessária para a captação desses elementos tem sido negligenciada, quer seja conscientemente pela limitação tecnológica ou financeira para a realização da pesquisa, quer seja inconscientemente pelo desconhecimento da importância e do papel desses elementos na produção linguística. Para pesquisas desse tipo seriam necessárias a preparação de um estúdio de gravação em que diferentes câmeras pudessem focalizar segmentos únicos da sinalização, o que não é comum com vídeos na rede. Ao pensarmos num corpus extraído da rede, no entanto, reduz-se a possibilidade de manuseio dessas variáveis²⁵. Os vídeos que apresentam limitadores de análise devem ser excluídos já no momento da seleção prévia, de

²⁵ Alguns programas *freeware* no entanto, podem ajudar na limpeza da qualidade do vídeo, como o aumento e diminuição da claridade, contraste, saturação, redimensionamento, rotacionamento, inversão, além de muitos outros ajustes de vídeo que serão bem úteis ao pesquisador. Programas como o windows live movie maker ou o virtual dub disponível em <http://www.virtualdub.org/>, são exemplos de programas que auxiliam nesse processo.

acordo com a finalidade da pesquisa, critério que será analisado mais detalhadamente na seção específica de seleção de vídeos do youtube.

3.4 Seleção dos vídeos

Definidos os métodos de extração, a plataforma que servirá como fonte de dados, a forma de organização do corpus no computador que servirá como armazenador do corpus é necessário começar a seleção dos vídeos. Como mencionado anteriormente, o corpus será definido sobretudo pela finalidade da pesquisa. A representatividade através dos principais gêneros textuais dispostos na rede foi o objetivo do corpus compilado aqui. Para isso, no entanto, propomos uma definição do conceito de gêneros discursivos e um levantamento de como esses gêneros são tratados em diversos suportes.

3.4.1 A questão dos gêneros discursivos

Desenvolver um corpus que pretenda agrupar textos dentro de uma determinada tipologia demanda algumas definições prévias. Primeiramente, é necessário entender que não se trata de um simples exercício de classificação geral ou parcial de gêneros. Enquanto materialização de atividades sócio-discursivas isso seria demasiadamente laboroso, senão impossível. Também não se justificaria quanto à relevância prática da pesquisa.

Entender que o uso da língua só ocorre por meio de gêneros discursivos, implica considerar que o uso, ensino e prática dos gêneros discursivos são elementos fundamentais para o estudo da linguagem.

Mapear, estudar e pesquisar a formação de um corpus de gênero a partir de dados da rede permite-nos desenvolver instrumentais linguísticos que orientem os pesquisadores de língua de sinais ao desenvolvimento de atividades fundamentais de pesquisa linguística. Possibilita-se não só o entendimento, mas o ensino, a aprendizagem e o uso de formas adequadas de realizações linguísticas em situações sociais particulares.

Entendemos que gêneros discursivos, conforme Bakhtin (2000, pág 279) trata-se de “tipos relativamente estáveis de um enunciado”.

Uma definição mais completa aparece no início do capítulo da obra dedicado ao tema:

O emprego da língua efetua-se em forma de enunciados (orais e escritos) concretos e únicos, proferidos pelos integrantes desse ou daquele campo da atividade humana. Esses enunciados refletem as condições específicas e as finalidades de cada referido campo não só por seu conteúdo (temático) e pelo estilo da linguagem, ou seja, pela seleção dos recursos lexicais, fraseológicos e gramaticais da língua mas, acima de tudo, por sua construção composicional. Todos esses três elementos – o conteúdo temático, o estilo, a construção composicional – estão indissolivelmente ligados no todo do enunciado e são igualmente determinados pela especificidade de um determinado campo da comunicação. Evidentemente, cada enunciado particular é individual, mas cada campo de utilização da língua elabora seus tipos relativamente estáveis de enunciados, os quais denominamos gêneros do discurso.

Dentro da abordagem em que trabalhamos, em que a língua estrutura-se a partir de finalidades sociais e interacionais, em que se privilegia a sua natureza funcional e interativa, considerando que é impossível usar a língua sem que seja através de um gênero do discurso, pode parecer impossível considerar a possibilidade de definição de gêneros visto que são infinitas as possibilidades e campos de atuação humano. Como o próprio Bakhtin pontua, a heterogeneidade funcional, poderia tornar os traços gerais dos gêneros discursivos demasiadamente abstratos e vazios. No entanto a própria definição de gêneros e o reconhecimento da relativa estabilidade atestam que é uma empreitada possível. Não só possível, como fundamental para o estudo da linguagem, como pontua ainda Bakhtin(2000, pág 279):

O desconhecimento da natureza do enunciado e a relação diferente com as peculiaridades das diversidades de gênero do discurso em qualquer campo da investigação linguística redundam em formalismo e em uma abstração exagerada, deformam a historicidade da investigação, debilitam as relações da língua com a vida.

Sabemos que Bakhtin não elabora uma tipologia de gêneros, mas os classifica em gêneros primários e secundários. Aqui temos dois pontos importante de atenção para essa e outras pesquisas que utilizam a rede como corpus de línguas de sinais. Em se tratando de línguas ágrafas, sobretudo de línguas de sinais, muitas com gênese tão recente como a língua de sinais da Nicarágua, não é tão simples classificar e distinguir esses gêneros. A definição de gêneros secundários de Bakhtin evidencia melhor essa complexidade: “os gêneros discursivos secundários surgem nas condições de um convívio cultural mais complexo e relativamente muito desenvolvido e organizado (predominantemente o escrito) – artístico, científico, sociopolítico, etc”. Desenvolvimento prescinde evolução cronológica, o que é algo difícil para línguas subalternizadas e até mesmo proibidas, como foram as línguas de sinais. Ao enfatizar a modalidade escrita, o autor ainda demonstra a necessidade de planejamento, elaboração, reflexão na produção, elementos que historicamente mostram-se somente passíveis de realização com o auxílio de ferramentas tecnológicas. Sem tecnologia que atenda à demanda de registro de línguas de sinais, esses elementos seriam renegados e a elaboração de enunciados secundários, na concepção de Bakhtin ficaria praticamente impossível.

O segundo ponto a ser considerado, quando tratamos da classificação Bakhitiniana de gêneros, é aquele que define os gêneros primários pelas condições de comunicação discursiva imediata. Sabemos que o elemento suporte também é definidor do gênero, o que possibilita que um mesmo texto, em suportes diferentes possa ser definido como pertencente a gêneros diferentes. Logo, a influência de estarmos buscando gêneros diferenciados em um suporte único deve ser considerada. Haverá elementos de suporte, nesse caso, que influenciarão a estrutura de todos os gêneros como por exemplo o grau elevado de preparação, performance e artificialidade em vídeos postados para visualização pública, como já mencionado. Essa necessidade auxiliaria a produção e o reco-

nhecimento de gêneros secundários mas afetaria o estudo de gêneros primários como a réplica do diálogo cotidiano.

Essa dificuldade também é recorrente na elaboração de corpus que não utilizam a rede, visto que o constrangimento e a consciência do registro sempre afeta de alguma forma a produção verbal dos participantes. Um paradoxo no entanto, é que na rede, mais do que em qualquer outro meio, é possível encontrar milhares de registros espontâneos, e que, portanto, representaria o grau máximo de naturalidade exigido e característico de um gênero primário. O fato é que a elaboração de um corpus desse tipo deve considerar os elementos linguísticos e também elementos de ordem mais ampla como condições específicas e finalidades do referido campo em que o enunciado foi formado. Isso implica buscar e caracterizar no corpus toda a motivação e cenário de sua formação. Aqui está um ponto que parece ser crítico na utilização dos dados da rede. Anonimato, informações escassas e inexistentes poderiam prejudicar a análise, não fosse a utilização de recursos que buscam suprir essas falhas, como observaremos na compilação do corpus.

Esses recursos envolvem um rigoroso critério de seleção, confirmação e validação dos dados, com o uso de informantes e do conhecimento prévio, linguístico e da comunidade pesquisada por parte do pesquisador. O ponto crítico aí não invalidaria a utilidade e importância da análise dos elementos linguísticos. Como menciona Geraldí (2010, pág. 72):

Isto não quer dizer, obviamente, que os recursos linguísticos deixem de ter qualquer influência: eles estão presentes e são responsáveis pelo acionamento e agenciamento do que lhe é exterior; estas informações, por seu turno, se refletem nos elementos estritamente linguísticos, de modo que a significação reconhecida se reveste do tema e das considerações externas que fazem, de fato, o “sistema” funcionar.

No ideal teórico em que trabalhamos, portanto, os gêneros se definem por aspectos sociocomunicativos e funcionais mas a estabilidade que eles produzem, refletida nos aspectos estruturais linguísticos, também são elementos importantes para a sua caracterização. Considerando, então, todos os elementos necessários para a compilação do cor-

pus, cabe fazer algumas considerações que esclareçam as definições de gênero do discurso.

Conforme Marcuschi (2010), quando nos referimos a aspectos lexicais, sintáticos, tempos verbais, relações lógicas, ou seja, aspectos estritamente linguísticos de uma sequência comunicativa, e usamos esses aspectos para a definição de uma sequência, estamos nos referindo a uma tipologia textual, predominando, portanto a identificação de sequências linguísticas típicas como norteadoras. Assim, em geral, os textos são tipologicamente variados, heterogêneos.

Logo, um texto em libras retirado da rede, pode conter inúmeras sequências tipológicas que não irão necessariamente definir o seu gênero. O texto é nomeado tipologicamente como “narrativo”, “descritivo”, “argumentativo” de acordo com a sua sequência de base, essa nomeação não corresponde ao gênero textual. Essa definição é importante, pois ao levantarmos trabalhos de corpora de língua de sinais e línguas orais na internet, percebemos que em alguns há uma classificação tipológica dos textos, enquanto em outros uma definição mais restrita.

O gênero textual se define pela ação em que ele aparece, é definido pela sua função, intenção, sem no entanto desconsiderar o papel relevante que as estruturas linguísticas operam na sua definição. O quadro sinótico de Marcuschi (2010) nos permite visualizar essas definições de forma mais clara:

TIPOS TEXTUAIS

1. constructos teóricos definidos por propriedades lingüísticas intrínsecas;

2. constituem seqüências lingüísticas ou seqüências de enunciados e não são textos empíricos

3. sua nomeação abrange um conjunto limitado de categorias teóricas determinadas por aspectos lexicais, sintáticos, relações lógicas, tempo verbal;

4. designações teóricas dos tipos: narração, argumentação, descrição, injunção e exposição

GÊNEROS TEXTUAIS

1. realizações lingüísticas concretas definidas por propriedades sócio-comunicativas;

2. constituem textos empiricamente realizados cumprindo funções em situações comunicativas;

3. sua nomeação abrange um conjunto aberto e praticamente ilimitado de designações concretas determinadas pelo canal, estilo, conteúdo, composição e função;

4. exemplos de gêneros: telefonema, sermão, carta comercial, carta pessoal, romance, bilhete, aula expositiva, reunião de condomínio, horóscopo, receita culinária, bula de remédio, lista de compras, cardápio, instruções de uso, outdoor, inquérito policial, resenha, edital de concurso, piada, conversação espontânea, conferência, carta eletrônica, bate-papo virtual, aulas virtuais etc.

Fonte: Quadro sinóptico de MARCUSCHI, L. A. **Gêneros Textuais**. Definições e Funcionalidade. Em *Gêneros Textuais e ensino*. Dionísio, A.P., Machado, A.R. e Bezerra, M.A.(org.) 5 ed. Rio de Janeiro. Lucerna. 2007

Definida a noção de gênero que iremos trabalhar, convém entendermos um pouco como os gêneros vem sendo compilados em corpus específicos.

3.4.1.1 Os gêneros em corpora

Definida a noção de gênero que iremos trabalhar, precisamos determinar quais os gêneros que buscaremos compilar no corpus-web. Para a definição dos gêneros da presente pesquisa propomos uma matriz de dados elaborada a partir dos seguintes levantamentos: levantamento dos gêneros apresentados em dois importantes corpora da língua portuguesa; nos principais corpora de língua de sinais do levantamento realizado por Konrad (2011); na classificação de gêneros do protocolo padrão de Metadados - ISLE Meta Data Initiative²⁶ (IMDI); em levantamento feito com a comunidade surda; em levantamento de pesquisas realizadas com gêneros específicos da língua de sinais brasileira a partir da plataforma de teses e dissertações de domínio público e, por fim, a partir de levantamento de amostragem de vídeos postados na plataforma que servirá como fonte, youtube.com.

3.4.1.1.1 Gêneros em corpora da língua portuguesa

O corpus do português²⁷ (Davies, Mark e Michael Ferreira, 2006), que se propõe um corpus representativo abrangente, utiliza apenas quatro grandes categorias para o que chama de gênero: acadêmicos, notícias, ficção e oral.

²⁶ Padrão de metadados criado para descrever recursos de multimídia e de linguagem multi-modal. Disponível em <http://www.mpi.nl/imdi/>

²⁷ Corpus com 45 milhões de palavras de quase 57,000 textos em português do século XIV ao século XX, disponível on-line em <http://www.corpusdoportugues.org>

Fig 16 - Interface de refinamento de busca do corpus do português de Davies, Mark e Ferreira

The screenshot shows the 'CORPUS DO PORTUGUÊS' search interface. At the top, it displays '45.000.000 PALAVRAS, s.XIII XX'. Below this is a search bar and a table of results. The table has columns for 'CONTEXTO', 'TOCOS', and frequency counts for years 1300s, 1400s, 1500s, 1600s, 1700s, 1800s, 1900s, and 'FREQ'. There are also columns for 'ACRQ', 'NTAS', 'FREQ', and 'OBS'. The table lists words like 'BELO', 'BELA', 'BONITA', 'BONITO', 'REBRETO', 'REBRETO', 'REBRAS', 'AGRADÁVEL', 'LINDO', 'ROMANOS', 'BRILLOS', and 'BONITAS'. The interface includes a search bar at the top right and a 'PESQUISA POR REGIÃO' button at the bottom.

Fonte: DAVIES, MARK and MICHAEL FERREIRA. (2006-) Corpus do Português: 45 million words, 1300s-1900s. Disponível online em <<http://www.corpusdoportugues.org>>

Os textos acadêmicos são aqueles retirados das fontes Enciclopédia Uol, com textos formatados e didaticamente escritos para alunos da educação básica e do site do Núcleo Interinstitucional de Linguística Computacional, com artigos científicos de nível superior. Por ser uma enciclopédia eletrônica, o textos da enciclopédia Uol²⁸ apresentam uma tipologia textual dissertativa-expositiva, semelhante a uma enciclopédia física, já que o site utilizado não apresenta recursos de hipertextualidade como referências a vídeos ou outros textos dentro do texto principal. A mesma tipologia textual predomina nos artigos científicos, porém numa vertente mais argumentativa do que expositiva. Nesse tipo de gênero o autor apresenta informações com o intuito de convencer o leitor de determinada concepção científica.

Os textos classificados como de ficção são formados por livros de ficção de diversos gêneros como drama, romance, contos, novelas e suspense. Obras de Euclides da Cunha, Monteiro Lobato, Jorge Amado e Érico Veríssimo foram incluídas nesse domínio. Os textos foram escaneados ou retirados da biblioteca digital de literatura da UFSC²⁹. Outros

²⁸ Os links remetem a uma versão da enciclopédia disponível somente para assinantes. Uma visualização de textos semelhantes pode ser obtida na versão biblioteca do mesmo site, disponível em <http://educacao.uol.com.br/biblioteca/>

²⁹ Os links não estão mais disponíveis. O link atual da biblioteca é <http://www.literaturabrasileira.ufsc.br/>

10 livros de ficção foram retirados do corpus lácio-web³⁰, corpus do português brasileiro escrito contemporâneo.

O corpus agrupou na classificação “oral” dois gêneros: notícias de jornal e entrevistas. As notícias foram retiradas de sites de notícias como uol, gazeta do povo, correio do povo, diário de Pernambuco, estado, dentre outros. Um exemplo de entrevista utilizada como fonte da classificação “oral” nesse corpus pode ser visto a seguir:

Fig 17 - Entrevista on-line utilizada pelo autor como fonte oral

Everaldo Lima d'Alverga Entrevista Junho 2001

Clóvis Scarpino
 Jornalista e Fotógrafo

Uma Lição de Vida

"A fotografia é o que o olho do cara vê, dentro de uma alimentação de luz e angulação e o cara tem que abrir a mente e a lente".



Aos 71 anos, 50 dos quais dedicado a profissão, **Clóvis Scarpino**, formado em Economia, Jornalista, documentarista, Fotógrafo e pesquisador do Samba, iniciou sua trajetória jornalística como repórter de rua, mas logo se rendeu aos encantos da fotografia. Em suas andanças pelo mundo do samba, conheceu Pixinguinha, Nelson Cavaquinho, Zé Kéti e Silas de Oliveira, conviveu com Ismael Silva e muitos bambas do samba e da imagem. Com exclusividade, concedeu ao também Jornalista e Fotógrafo **Everaldo d'Alverga**, uma Entrevista onde o “papo” foi marcado pela total descontração e naturalidade, trazendo-nos fatos, curiosidades e acontecimentos que marcaram a história do jornalismo em nosso País, numa verdadeira aula de vida.

EVIRT: Não poderia deixar de fazer a pergunta óbvia e básica. Como é que você começou?

SCARPINO: Comecei no Jornal Metropolitano, jornal da UME, União Metropolitana do Estudante, que saía aos domingos no encarte do Diário de Notícias, aí eu comecei lá.

Fonte: <http://www.evirt.com.br/entrevista/scarpino.htm>

Há ainda uma listagem de fontes orais, provavelmente gravadas, de linguagem falada no Recife e de gravações “orais” sem mencionar a fonte. Os autores disponibilizam uma lista das fontes utilizadas com os títulos dos livros e link’s web para as fontes eletrônicas.

³⁰ Disponível em <http://www.nilc.icmc.usp.br/lacioweb/index.htm>

Fig 18 - Visualização da lista de fontes do corpus do português

	A	B	C	D	E	F	G
1	lexID	genre	# palav	titulo	autor	data	fonte
2	26214	Bras-Fic	41742	Beltrão, Luiz	A Greve dos Desempregados	1984	SCANNED
3	26224	Bras-Fic	39392	Comparato, Doc	A Guerra das Imaginações	1997	SCANNED
4	26216	Bras-Fic	40579	Callado, Antônio	A Madona de Cedro	1957	SCANNED
5	26260	Bras-Fic	40465	Vieira, José Geraldo	A Mais que Branca	1974	SCANNED
6	26248	Bras-Fic	42339	Penna, Cornelio	A Menina Morta	1958	SCANNED
7	26249	Bras-Fic	42091	Queiroz, Dinah Silveira de	A Muralha	1954	SCANNED
8	26239	Bras-Fic	39704	Montello, Josué	A Noite sobre Alcântara	1978	SCANNED
9	26226	Bras-Fic	42225	Costa, Eduardo Alves da	A Sala do Jogo	1989	SCANNED
10	26052	Port-Fic	2243	A CONSOADA	ABEL BOTELHO		SCANNED
11	26006	Port-Fic	2526	Trânsito Impedido	AFONSO BOTELHO		SCANNED
12	26003	Port-Fic	2471	Hora de Folga	AFONSO RIBEIRO		SCANNED
13	26122	Port-Fic	57165	O homem sem nome	Aguar, João	1986	SCANNED
14	26117	Port-Fic	43589	Os incuráveis	Agustina Bessa Luis	1982	SCANNED
15	26015	Port-Fic	2405	O Noivo	AGUSTINA BESSA LUIS		SCANNED
16	26120	Port-Fic	43225	Margem Norte	Alexandre Cabral	1979	SCANNED
17	26059	Port-Fic	1826	O CAGADO	ALMADA-NEGREIROS		SCANNED
18	26026	Port-Fic	2443	PONITA TENETE	ÁLVARO GUERRA		SCANNED
19	26284	Bras-Fic	42688	Memorial de um Passageiro de B	Amadeu Amaral	1921	http://alecrim.inf.ufsc.br/bdnpull/
20	26285	Bras-Fic	2749	Novela e Conto	Amadeu Amaral		http://alecrim.inf.ufsc.br/bdnpull/
21	26137	Port-Fic	38529	A Máscara e o Destino	Amonim, Antonio Guedes de	1944	SCANNED
22	25662	Bras-Acad	28208	O ensaio do conceito de tempo: c	André Ferrer Pinto Martins		http://www.nilc.icmc.usp.br/facioweb/corpora.htm
23	25669	Bras-Acad	15906	Análise de redes metabólicas em	Andreas Karoly Gombert		http://www.nilc.icmc.usp.br/facioweb/corpora.htm
24	25684	Bras-Acad	18515	Uso de Redes Neurais para Corre	Andréia Gentil Bonfante		http://www.nilc.icmc.usp.br/facioweb/corpora.htm
25	26119	Port-Fic	40757	Os nos e os Laços	Antonio Alçada Baptista	1985	SCANNED
26	26077	Port-Fic	2462	Os Dons Barbeiros	ANTONIO ARNAUT		SCANNED
27	26278	Bras-Fic	12484	Contos Avulsos	Antônio Castilho de Alcântara Machado		http://alecrim.inf.ufsc.br/bdnpull/
28	26280	Bras-Fic	10398	Brás, Bexiga e Barra Funda	Antônio Castilho de Alcântara Machac	1927	http://alecrim.inf.ufsc.br/bdnpull/
29	26281	Bras-Fic	14871	Laranja-da-China	Antônio Castilho de Alcântara Machac	1928	http://alecrim.inf.ufsc.br/bdnpull/
30	26282	Bras-Fic	17820	Mana Maria	Antônio Castilho de Alcântara Machac	1935	http://alecrim.inf.ufsc.br/bdnpull/

Fonte: <<http://www.corpusdoportugues.org/help/cdp.xls>>

Por fim, novamente fontes como sites de notícias como uol, gazeta do povo, correio do povo, diário de Pernambuco, estadão, estado dentre outros do tipo, foram utilizados para compor a classificação “notícia”, que aparece na listagem como “periódicos”. Também foram utilizados como fonte, notícias disponíveis em banco disponibilizado pelo projeto Linguateca³¹, um centro de recursos para o processamento computacional da língua portuguesa.

Outro grande projeto, o projeto AC/DC³² - O Corpus Brasileiro é uma coletânea de aproximadamente um bilhão de palavras do português brasileiro, resultado de projeto coordenado por Berber Sardinha, (GELC, LAEL, Cepril, PUCSP), com financiamento da Fapesp. O projeto classifica como gênero, 13 domínios discursivos, subdivididos em 25 fontes, conforme está disposto no quadro a seguir:

Classe	Gênero	Fonte
ei	Acadêmico	Artigos

³¹ Disponível em <http://www.linguateca.pt/>

³² Disponível em <http://www.linguateca.pt/acdc/>

ej	Acadêmico	Teses e dissertações
eq	Acadêmico	Anais de congresso
en	Cinema e TV	Roteiros
ee/ef	Educação	Diversos
et	Enciclopédia	Wikipédia
fa	Esporte	Narração de jogos de futebol
ek	Informática	Manuais
ed	Jornalismo	Revistas
eg	Jornalismo	Jornais
eo	Jornalismo	Horóscopo
fe	Jornalismo	Entrevistas
em/ex	Legislação	Diversos
eb	Literatura	Contos
ec	Literatura	Crônicas
eh/ew	Literatura	Variados
eu	Literatura	Biografias
ea	Medicina	Bulas de remédio
ep	Política	Atas de assembléia legislativa
fb	Política	Debates de TV
fc	Política	Pronunciamentos do presidente
fd	Política	Sessões do congresso
el	Religião	Diversos
ev	Religião	Bíblia
er/es	Técnico	Relatórios e manuais diversos

Quadro de domínios discursivos do corpus AC-DC disponível em
<http://www.linguateca.pt/acdc/>

Cabe destacar, no entanto que 20 dessas fontes/registros referem-se ao português escrito e apenas 5 referem-se ao português falado.

As amostras aparentam denotar uma escolha determinada mais pela disponibilidade dos textos online e acesso aos mesmos do que pela relevância do gênero. Tal procedimento, verificado em corpus da língua portuguesa, justifica a relevância de um levantamento prévio na plataforma *youtube* para verificação dos gêneros disponíveis.

3.4.1.1.2 Gêneros em corpora de línguas de sinais

Como já mencionado, o gênero escolhido e a fonte do corpus serão diretamente determinados pela finalidade da pesquisa. Utilizamos o levantamento feito por Reiner Konrad (2011), que listou 34 projetos que ele classificou como projetos de corpus de línguas de sinais e a partir dele selecionamos os projetos que tinham como objetivo constituir uma documentação linguística abrangente da língua de sinais referida. Vejamos que tipos de gêneros são selecionados nesses projetos:

Corpus	Total de horas de filmagens	Gêneros	Endereço na rede
DGS Sign Language Corpus	1159 Horas	Diálogos Monólogos Conversação Livre Narrativas eliciadas Ítems lexicais eliciados	http://www.sign-lang.uni-hamburg.de/dgs-korpus/
Projeto DGS.txt.ty p de tipos de texto da língua gestual alemã	620h	-Conversação livre -Diálogos -Narrativas -Relatórios -Contos de Fadas -Histórias infantis -Piadas -Poemas -Textos Informativos -Debates -Entrevistas -Sermões	http://www.sgw.hs-magde-burg.de/dgstxttyp/index2.html
Corpus da Língua de Sinais	71h30	- Diálogos - Narrativas	http://www.ru.nl/corpus_ngt/

Holandesa			
Corpus da Língua de Sinais	106h	- Diálogos - Entrevistas	http://www.umar7023.cnrs.fr/Objectifs-du-projet.html
Corpus da Língua de Sinais Britânica	180h em crescimento	- Narrativas de experiências pessoais - Conversações livres - Entrevistas - Palavras eliciadas - Sentenças eliciadas - Narrativas eliciadas	http://www.bslcorpusproject.org/
Língua de Sinais Australiana	300h	- Conversação em grupo - Entrevistas - Narrativas eliciadas - Palavras eliciadas	http://www.auslan.org.au/about/corpus/
Língua de sinais da comunidade indígena Kata Kolok – Bali	100h	- Diálogos - Monólogos - Narrativas culturais	http://www.uclan.ac.uk/rESE-arch/environment/groups/islands.php
Língua de Sinais do Mali	30h	- Diálogos - Monólogos - Discussões - Conversação livre	http://www.hum.leiden.edu/lucI/research/past-projects/lasima.html

Considerando a noção Bakhtiniana da complexidade de evolução da língua mediada principalmente por tecnologias de registro necessárias para a produção de gêneros discursivos secundários, torna-se claro a preferência por gêneros primários na formação dos corpora listados.

A recorrência pode ser listada da seguinte forma:

Gênero/Tipos	Presente nos corpus
Diálogos/Conversas	100%
Narrativas	87,5%

Entrevistas	50%
-------------	-----

De fato, conforme apontado por Leite (pág.134) em alusão as afirmações de Schegloff (1996:54) e Clark(2000: 54) a interação face-a-face além de ser um cenário primordial da vida social e gênero de fala universal, constitui-se no berço para a aquisição de linguagem.

Paradoxalmente, os procedimentos metodológicos técnicos e específicos para a captação desse gênero são os mais complexos e, portanto, incomuns em plataformas da rede que se caracterizam principalmente pela produção caseira e sem aparato tecnológico profissional.

3.4.1.1.3 Gêneros em padrão de metadados

O instituto Max Planck de Psicolinguística em Nijmegen³³, na Holanda, desenvolveu um arquivo de dados para armazenar grande quantidade de material linguístico exclusivo para uma grande variedade de idiomas em todo o mundo que é registrado e analisado por pesquisadores de diversas áreas linguísticas. Para navegar por esses dados foi criado o ISLE Meta Data Initiative, um padrão de metadados para descrever recursos de multimídia e de linguagem multi-modal. A escolha desse protocolo foi motivada pois ele inclui o projeto ECHO – Sign language³⁴, que contém corpus de quatro línguas européias de sinais: Holandesa, Britânica, Sueca e Alemã e línguas ameaçadas. Além disso, o projeto concentra-se na descrição e armazenamento de línguas ameaçadas, o que sugere uma adequação a gêneros de línguas ágrafas e em situações próximas às das línguas de sinais.

Para classificação de gêneros o padrão de metadados divide os textos em gêneros e sub-gêneros. As categorias principais de gênero são:

³³ Disponível em <http://www.mpi.nl/>

³⁴ Disponível em http://sign.let.ru.nl/groups/slcwikigroup/wiki/4a7fb/ECHO_project.html

Documentos secundários

Conteúdo que se refere, ou comenta, partes de dados primários. Ex. Uma gramática, uma resenha de um livro.

Literatura

Conteúdo que narra um evento imaginário e é valorizado pela beleza estética da língua

Ex. Um pequeno romance, um drama

Poesia

O Conteúdo está disposto em forma de versos ou padrão similar. Ex. Uma balada, uma epopéia

Canções

O conteúdo é executado para uma melodia. Ex. Uma canção popular, uma canção de ninar

Ficção popular (Folclore)

O conteúdo narra um evento imaginário que apela para o gosto popular. Exemplos: um romance policial, uma história de ficção científica.

Textos de ritual ou religiosos

O conteúdo diz respeito ao desempenho de ritos religiosos que consistem em tipos de discurso prescritos. Exemplos: uma oração, um ritual de cura, um catecismo

Artigos de jornal

O conteúdo é não-ficcional distribuído por meio de um jornal, uma revista ou pela internet. Exemplos: um ensaio político, um relatório científico.

Material de TV/Rádio

O conteúdo é um texto falado /sinalizado não-ficcional que é transmitido via TV, rádio ou internet. Exemplos: uma discussão política, um documentário sobre a vida animal.

Discurso

O conteúdo é composto por declarações proferidas /sinalizadas por um ou mais indivíduos. Eles são produzidos com o objetivo de comunicar um pensamento ou intenção para os interlocutores presentes ao evento. Exemplos: uma lenda, uma conversa, um discurso público.

Teatro/Drama

O conteúdo é uma peça de ficção que é atendido no palco ou para a radiodifusão. Exemplos: um filme, uma peça de teatro, uma leitura pública de um livro.

Anotações pessoais

O conteúdo é um breve registro de fatos ou pensamentos que atuam como uma ajuda de memória. Exemplos: notas de campo, notas para um discurso público.

A classificação do protocolo congrega vários elementos para a definição de categorias estáveis. Ora o suporte é considerado como definidor da categoria (Material de rádio e tv), ora o domínio (texto de ritual ou religiosos), ora a função (anotações pessoais). Consideraremos os gêneros como os exemplos dados em cada categoria que apresentam relativas autonomias de função, estilo, conteúdo e composição, sem considerar os aspectos relativos ao canal utilizado.

3.4.1.1.4 Gênero nas demandas da comunidade surda

Outro indicador para a definição de categorias de gênero do corpus pode vir da própria demanda apresentada pela comunidade de usuários da língua pesquisada. Um questionário, elaborado com a finalidade de levantar parte dessas demandas foi produzido em língua portuguesa e apresentado em língua de sinais a 10 acadêmicos surdos. O formulário entregue para os entrevistados não continha a pergunta em português, apenas espaços para as respostas. Para alguns dos entrevistados o questionário foi apresentado pessoalmente e para outros online, através de vídeos gravados e postados na plataforma youtube. O levantamento apontou algumas questões importantes quanto à elaboração do corpus, dentre elas: necessidade de uma maior participação da comunidade surda nos projetos linguísticos desenvolvidos no Brasil; preocupação da comunidade com o registro da língua especialmente ao seu estado considerado “nativo” ou “puro” que estaria sendo influenciado pelo português; preocupação com um ensino da língua de sinais que efetivamente é utilizada pela comunidade surda, refletindo na formação de intérpretes e sinalizantes de qualidade.

Quanto aos gêneros relevantes para um corpus de amostragem os principais indicados pelos entrevistados por ordem de relevância foram: narrativas surdas (incluindo aqui entrevistas, narrativas de histórias pessoais e de sua comunidade), contos de histórias infantis, piadas surdas, literatura surda (poesia) e bate papo entre surdos.

3.4.1.1.5 Gênero na pesquisa linguística da libras

Outro indicador importante para a constituição do corpus foi as pesquisas científicas na área da linguística da língua de sinais que vêm sendo desenvolvidas nos últimos dez anos e que apontam para elementos marcadores de gêneros textuais em libras. Esse levantamento também permite análises posteriores dos dados do corpus quanto às discussões de características funcionais e estruturais de cada gênero.

Como critério para a realização do levantamento utilizaram-se os seguintes elementos: pesquisas de mestrado e doutorado que se concentram na grande área linguística, linguística histórica e linguística aplicada e que direta ou indiretamente analisasse aspectos de um gênero textual em libras.

Utilizamos a plataforma governamental biblioteca digital brasileira de teses e dissertações³⁵ como fonte para o levantamento. Os seguintes dados resultaram do levantamento:

Título	Pesquisa	Gênero textual	Observação
A construção da argumentação na língua brasileira de sinais: divergência e convergência com a língua portuguesa	Tese de Doutorado – UFPB – Linguística http://www.dominiopublico.gov.br/pesquisa/detalheobraform.do?select_action=&co_o	Conversação espontânea e semi natural	Apesar de trabalhar com foco na argumentação a autora faz consideração sobre alguns aspectos observados na conversação natural em um corpus de 12 filmagens de crianças surdas con-

³⁵ Disponível em <http://bdtd.ibict.br/>

	bra=199735		versando com seus pares em ambiente escolar e em duas ocasiões em atividade eliciada. Embora tenha selecionado apenas participantes filhos de pais ouvintes e com aquisição de libras tardia, o tempo médio de utilização da língua de sinais é de 5 anos
Investigando a categoria aspectual na aquisição da língua brasileira de sinais	Dissertação de mestrado – UFSC - Linguística http://www.dominiopublico.gov.br/pesquisa/detalheobraform.do?select_acion=&co_obra=190365	Conversação espontânea em período de aquisição de linguagem	A autora analisa a interação de uma criança surda, filha de pais surdos e com uma irmã surda. São observadas situações de interação espontânea do dia a dia em 32 sessões de 30 a 40 minutos em um período de 3 anos.
Uma descrição do processo de referência em narrativas contadas em língua de sinais brasileira (libras)	Dissertação de mestrado http://www.teses.usp.br/teses/disponiveis/8/8139/tde-06052013-112529/pt-br.php	Narrativas	Apesar de analisar especificamente o processo de referência há elementos suficientes na pesquisa para apontar algumas características da narrativa em libras
Literatura surda: produções culturais de surdos em língua de sinais	Dissertação de Mestrado http://www.lume.ufrgs.br/handle/10183/32311	Literatura Surda	O foco da pesquisa é a análise da forma como os surdos vêm apresentando e construindo a Literatura Surda, com foco na língua de sinais
A segmentação da língua de	Tese de Douto-	Conversação	O pesquisador busca oferecer critérios para

sinais brasileira (libras): um estudo lingüístico descritivo a partir da conversação espontânea entre surdos	rado http://www.teses.usp.br/teses/disponiveis/8/8147/tde-25092008-160005/pt-br.php		a segmentação do discurso na língua de sinais brasileira (libras) em unidades gramaticais a partir de um corpus de conversação em libras
Indicadores de formalidade no gênero Mo nológico em libras	Dissertação de Mestrado de	Formalidade/Editais de concursos	O pesquisador busca apontar alguns indicadores de formalidade utilizando como corpus editais em libras

3.4.1.1.6 Gênero na plataforma de pesquisa

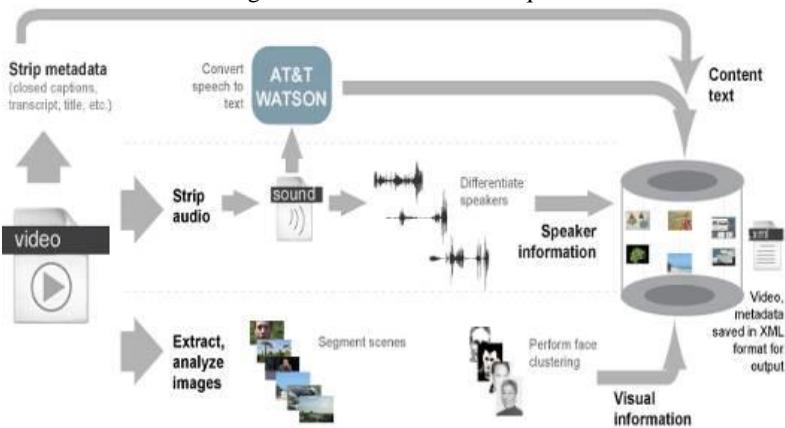
Entendendo que o próprio corpus pode indicar tendências de utilização e postagem de gêneros textuais na rede, propomos um levantamento que buscasse indicar padrões nos vídeos postados. Para isso é necessário primeiramente entendermos que informações podem ser incluídas sobre os vídeos postados e como elas são classificadas.

Metadados são informações sobre o vídeo como quem é o autor do vídeo, data de criação, duração etc. Na Internet costuma-se usar uma linguagem de computação chamada XML para codificar metadados. Nos vídeos, há dois tipos de metadados: aqueles que podem integrar o próprio código do vídeo e metadados externos a partir da página onde o vídeo está.

A seguir reproduzimos uma figura que retrata de forma simplificada os metadados que geralmente integram o vídeo, encontrada no projeto de pesquisa em metadados de um dos principais fornecedores mundiais de serviços de comunicação baseados em IP, a AT&T Researchers³⁶.

³⁶ Disponível em <http://www.research.att.com/>

Figura 19 - Metadados em arquivos de vídeo



Fonte:

<http://www.research.att.com/projects/Video/metadata.html?fbid=h9X8QrYyUdv>

Cada uma dessas informações do vídeo pode aparecer simultaneamente ou de forma isolada. Para arquivos multimídia ou conteúdos de vídeo os metadados podem ser representados de diversas formas. Cada forma determina a indexação do vídeo (capacidade de ser acessado rapidamente) por sistemas de recuperação de informação, como bancos de dados. Conforme vemos ainda na descrição do site:

- Metadados simples, contém informações básicas do vídeo (título, data, descrição, informações ator, e links de hipertexto para materiais relacionados).
- Conteúdo textual capturado de legendas, transcrições e legendas fechadas. Estas formas de conteúdo textual muitas vezes são os mais confiáveis, porque eles são criados manualmente por editores e provedores de conteúdo.
- Conteúdo textual criado automaticamente derivado da fala (diálogo e narração). O reconhecimento de voz é efetuado por vários softwares automáticos hoje em dia, como os da empresa privada Watson At & T indicado na figura acima.

- Conteúdo de informação visual calculado com técnicas de análise de vídeo que detectam mudanças no cenário (um fade, cortar, dissolver, etc) e detectam a face para encontrar personagens recorrentes ou atores em um vídeo.
- Segmentação de voz que permite a identificação de fala de diferentes pessoas no vídeo e facilita a outros processos automáticos, como síntese e reconhecimento de voz (a associação automática de um rosto com uma voz).

Os metadados inseridos pelo postador configuram as primeiras informações que serão utilizadas pelo pesquisador para organização do corpus. Elas basicamente se constituem de título, descrição, palavras-chave (tags), configurações de privacidade e categoria. Essas categorias se remetem a uma classificação prévia da temática do conteúdo do vídeo. Essa classificação é dividida nas categorias animais, automóveis, ciência e tecnologia, comédia, educação, entretenimento, esportes, games e desenhos, Guias e estilo, Jogos, Música, Notícias e política, Pessoas e blogs, Sem fins lucrativos/ativismo, Viagens e eventos

Caso deseje, o postador ainda poderá inserir informações avançadas, que se constituem de configurações sobre a possibilidade de perguntas e respostas no vídeo, sobre a licença e distribuição do vídeo, local e data em que foi gravado, configurações para vídeo 3d e estatísticas do vídeo. Nenhuma dessas informações é obrigatória. Com apenas dois cliques o usuário pode enviar o vídeo sem a necessidade de incluir metadados adicionais. Para fazer a busca de vídeos na plataforma é necessário apenas inserir uma entrada no campo principal de busca. Esse tipo de pesquisa é a forma mais ampla de pesquisa na plataforma youtube. Ela considera a palavra que aparece tanto no título quanto na descrição do vídeo que é inserida pelo postador.

Obviamente por se tratar de metadados de entrada postados pelos usuários, eles nem sempre correspondem às informações disponíveis nos vídeos, mas uma busca mais refinada tende a encontrar vídeos mais condizentes com a descrição.

O título do vídeo e a descrição podem indicar o propósito do vídeo, o autor, a condição linguística do autor (se surdo ou ouvinte), língua utilizada no vídeo, origem do personagem que fala no vídeo, local

de realização e variação sociolinguística utilizada. No vídeo a seguir, por exemplo, temos o contexto de produção do vídeo e o público a quem ele é destinado:

Fig 20 – Tela de informações e descrição de vídeo postado na plataforma youtube.com



Fonte: <http://www.youtube.com/watch?v=iqej4up7ogi>

O status linguístico do postador também pode ser encontrado de outras formas. O “nickname”, nome utilizado para o canal do postador do vídeo, já pode fornecer informações iniciais sobre o postador, se pessoa física ou organização e o tipo de canal, conforme vemos a seguir:

Fig 21 – Visualização da informações de canal específico do youtube



Fonte:

<https://www.youtube.com/watch?v=QtFO0-VODc>

Partindo das informações de localização de metadados nos vídeos e na rede, prosseguimos com a primeira busca geral na plataforma para que pudéssemos analisar os gêneros textuais presentes nos vídeos. Utilizamos como principal entrada da busca a referência mais comum a língua de sinais dos surdos no Brasil: Libras. O universo de vídeos que resultou da busca totalizou 467.000 vídeos. Seleccionamos os 100 primeiros vídeos e os classificamos a partir de sua categoria, descrição, língua utilizadas, status dos protagonistas dos vídeos e objetivo principal. O detalhamento do levantamento pode ser visto na tabela em anexo. A partir dos resultados classificamos os 100 primeiros vídeos apresentados.

Para classificação da língua principal utilizada considerou-se aquela que predominava no vídeo. Quando a utilização de outras línguas não era significativa, mas recorrente, explicitou-se seu uso. Foi observado em vídeos feitos em português oral mas que o orador utilizava exemplos ou apresentava trechos de sinais em libras durante o vídeo.

Para a pré definição do status linguístico dos oradores dos vídeos (se surdo ou ouvinte sinalizador) alguns critérios foram utilizados: (i) apresentação do orador no vídeo - na grande maioria dos vídeos os oradores se apresentam como surdos ou ouvintes no início do vídeo; (ii) a identificação do canal do vídeo - a grande maioria dos vídeos é postada pelo autor e proprietário do canal, clicando no link do canal encontramos os diversos vídeos do autor que o identificam, como vídeos na

escola de surdos que frequenta, vídeos em que fala de sua história, vídeos das associações a que está relacionado; (iii) uso da língua portuguesa como estrangeiro - marcas de falante de segunda língua expressas na descrição e título do vídeo; (iv) indicações de links para página no facebook, orkut ou blog - onde podem ser encontradas mais informações do usuário, inclusive seu status linguístico; (v) conhecimento linguístico do pesquisador quanto a marcas linguísticas apresentadas na produção de falantes de libras como segunda língua e por último, contato com o autor - para os vídeos selecionados para o corpus, contato direto com o autor do vídeo.

A seguir, descrevemos o percentual de classificação das línguas apresentadas nos vídeos:

63% - Libras (principal) e língua portuguesa em legenda e/ou oral sem imagem do orador ouvinte – simultâneas

15% - Somente Libras

12% - Língua Portuguesa oral (principal) com alguns exemplos em libras durante o discurso

2% - Libras com transcrição de glosas em português oral

5% - Língua portuguesa com janela de interpretação em libras

2% - Somente língua portuguesa

Quanto ao status dos oradores do vídeo:



O índice de falta de informação do status do sinalizante (6%) pode ser considerado baixo para a plataforma já que a inserção de

informações é opcional. A grande maioria dos vídeos contém falantes não nativos (ouvintes) sinalizando em libras. Apenas 14% dos vídeos são sinalizados por surdos. Dos 14 vídeos sinalizados por surdos, oito são destinados ao ensino de língua de sinais para ouvintes. Desses, três são feitos em forma de teatro de cenas do cotidiano para ensinar ouvintes. Apenas seis vídeos são destinados para surdos, sendo três do gênero piada, uma apresentação de negócio, e dois avisos de utilidade pública, um sobre uma manifestação política e um sobre orientações do que é *bullying*. Embora a simples utilização da entrada “libras” não permitisse afirmar inicialmente que essa seria a língua utilizada em todos vídeos, a análise mostrou que ela esteve presente em praticamente toda a amostra. Dos 100 vídeos selecionados, apenas o vídeo 63 não é realizado em língua de sinais. O Vídeo 67 também, embora apresente durante o vídeo algumas cenas do cotidiano de professores ensinando crianças surdas utilizando libras, apresenta apenas a língua portuguesa oral durante o vídeo. Os demais vídeos são constituídos ou por produções originalmente em libras ou por produções em língua portuguesa interpretada em libras, conforme apresentado no gráfico anterior.

Do levantamento feito para a entrada “libras”, 22 vídeos apresentaram classificação de categoria incompatível com o vídeo postado. Alguns vídeos permitem enquadramento em mais de uma categoria e a classificação escolhida pelo postador pode indicar a finalidade do vídeo postado.

Como os resultados da busca inicial, a partir dos dados apresentaram um número muito baixo de vídeos produzidos por surdos para surdos, apenas 6%, propomos uma nova busca na plataforma sem inserção de filtros, não mais a partir da língua, mas a partir da denominação mais comum para os usuários: Surdo.

Como resultados tivemos aproximadamente 37.500 vídeos com a palavra “surdo” em alguma parte do título ou descrição do vídeo. Uma análise dos 100 primeiros vídeos da busca apontou para um quadro diferenciado da busca por “libras”.

A análise da língua utilizada no vídeo apresentou os seguintes resultados:

Língua Utilizada no vídeo	
Somente Libras	24

Libras (principal) e língua portuguesa em legenda e/ou oral sem imagem do orador ouvinte – simultâneas	23
Libras (principal) e interprete junto oral	4
Libras com transcrição de glosas em português oral	2
Língua portuguesa com janela de interpretação em libras	6
Outras línguas de sinais	6
Outra língua oral com interpretação em outra língua de sinais	2
Língua Portuguesa oral (principal) com alguns exemplos em libras durante o discurso	12
Outra língua oral com tradução em português	2
Somente língua portuguesa	19

Enquanto para a primeira busca tivemos 15% de vídeos somente em libras, portanto prioritariamente destinado ao público surdo e de usuários de libras, para a busca “surdos” tivemos 24% dos vídeos realizados somente em libras. Apesar de um índice maior em material específico, no total, a quantidade de material acessível de alguma forma em libras foi menor que na primeira busca, 59% contra 85%. Nessa busca tivemos a particularidade de materiais produzidos e/ou traduzidos em outras línguas de sinais, 8%.

Quanto ao status dos oradores dos vídeos o levantamento consolidou-se da seguinte maneira:

Condição linguística dos oradores principais dos vídeos em relação a Libras	
Surdos utilizando libras	48
Ouvintes falantes de LP	29
Ouvintes falantes de outras línguas orais	4
Surdos e ouvintes sinalizadores	0
Deficiente auditivo/surdo oralizado	6
Ouvintes aprendizes	4
Ouvintes usuários da libras	4
Surdos de outros países	3
Apenas a escrita do português	1
Indeterminado	1

Portanto, embora a quantidade de vídeos produzidos ou traduzidos em libras tenha sido menor, a quantidade de falantes nativos nos vídeos com o resultado da busca “surdo” foi bem maior, 51% nessa busca, contra 16% na primeira. Enquanto tivemos 67% de ouvintes sinalizando na primeira busca, aqui tivemos 8%. Para a formação de um corpus com vídeos de falantes nativos da língua, esse tipo de busca parece indicar resultados mais adequados. Logo, para a montagem do corpus, sugere-se que a busca pelo status dos sinalizadores, *surdos*, seja incluída à busca pela língua, *libras*.

Analisando os 48 vídeos produzidos por surdos em língua de sinais brasileira quanto ao gênero textual, obtivemos os seguintes resultados:

Gênero	Descrição dos vídeos enquadrados no gênero	Número
Cotidiano	Apresentam cenas do cotidiano. Os autores dos vídeos filmam churrascos, intervalos na escola, eventos de família, futebol, crianças no seu cotidiano etc.	15
Artigo de opinião	Nesse tipo de vídeo o autor faz uma filmagem estática, em que a câmera é direcionada a uma única pessoa (geralmente o próprio autor do vídeo e dono do canal) em que o autor apresenta sua opinião sobre um assunto. Em geral de característica dissertativa-argumentativa o autor trata de um tema polêmico. O tema surge em geral por uma postagem iniciada por alguém sobre o assunto que dependendo da abordagem pode se desdobrar em dezenas de vídeos-resposta sobre o assunto. O autor estrutura o texto sobre argumentos científicos, heurísticos ou supostos, concluindo com sua opinião final sobre o tema, ou estimulando novos posicionamentos dos debatedores.	7
Piada	No formato estrutural tradicional de piada, estão vídeos em que os oradores narram uma estória fictícia (em geral em forma de monólogo) com um final inusitado que produz o efeito humorístico.	6
Humor	Aqui havia vídeos com o objetivo final de produzir efeito humorístico. Um dos vídeos apresentava uma piada encenada, em que dois atores	3

	<p>representam teatralmente as personagens da piada, outro vídeo de um canal administrado por um grupo de surdos em que os autores encenam desafios inusitados para produzir efeito humorístico como comer pimenta com leite, andar de cueca na rua etc e um último vídeo onde os surdos realizam entrevistas engraçadas com vários colegas na escola em que fazem piadas com as formas de sinalizar e os trejeitos dos colegas.</p>	
Cinema amador	<p>Vídeos em que pequenos filmes caseiros são produzidos, sobre as mais diversas temáticas, com o objetivo de produzir entretenimento para o público surdo e promover a atuação dos atores amadores. Tenta-se utilizar a linguagem cinematográfica, com a limitação do amadorismo.</p>	3
Entrevista	<p>Vídeos em que indivíduos surdos são entrevistados, ou por outros surdos, por ouvintes sinalizadores ou por ouvintes em língua portuguesa com auxílio de intérprete.</p>	3
Avisos/Informativos	<p>Aqui estão vídeos em que o objetivo é trazer informações úteis para os surdos. Marcado por uma linguagem objetiva, com pouca argumentação pessoal, o autor utiliza uma tipologia dissertativa-expositiva em que o objetivo principal é repassar uma informação. O foco não está na argumentação</p>	3
Teatro para web	<p>Surdos produzindo teatro sobre temas cotidianos. Predomínio da lin-</p>	2

	guagem teatral em que são cenas são reproduzidas sem uso de linguagem cinematográfica, como efeitos, mudanças de ângulos etc.	
Convites	Vídeos em que surdos convidam outros surdos para eventos específicos. Utilizam de linguagem objetiva.	2
Ensino Educação	Vídeo em que professores surdos ensinam sobre a história da educação de surdos	1
Pregação protestante	Vídeo em que um pastor surdo protestante prega em um culto religioso	1
Canção interpretada por surdos	Vídeo em que um coral de surdos, com o auxílio de uma regente ouvinte apresenta uma música	1
Trabalhos acadêmicos	Vídeo produzido por alunos surdos de nível superior como atividade prática de uma disciplina	1

Os gêneros mais recorrentes apresentados, portanto, foram: cotidiano, artigos de opinião, piada (monólogos), humor (encenações humorísticas), cinema amador, entrevistas, avisos/informativos, teatro e convite. Devido ao objetivo do vídeo, que é apresentar a cena cotidiana, os vídeos do gênero “cotidiano” não são planejados para uma boa apresentação do texto em língua de sinais. Em geral o conteúdo linguístico concorre com a apresentação da cena, personagens, performance etc. Mais da metade dos vídeos possuem mais tempo de filmagem do evento do que de informação linguística. Embora em alguns os indivíduos que são filmados se apresentem e façam algum tipo de comentário contextualizador da ocasião, na grande maioria são surpreendidos no meio da atividade, sem possibilidade de recuperação do contexto. A performance linguística não é pensada para o suporte da internet, dessa forma há cortes nos textos (a filmagem começa ou termina antes do final do discurso filmado), não há preocupação com iluminação, clareza das infor-

mações ou com os destinatários da web. Por esse motivo esses vídeos foram excluídos da análise.

Os gêneros humor (encenações), cinema amador e teatro, também foram excluídos da análise. Devido à natureza amadora, há uma forte artificialidade nos diálogos, além de uma nítida influência da estruturação de diálogos em língua portuguesa nos respectivos gêneros, possivelmente uma tentativa de reprodução de falas observadas em legendas, no caso dos programas humorísticos e do cinema. Logo, consideramos como indicadores de gêneros recorrentes no suporte, úteis para a definição dos gêneros a serem compilados na pesquisa, os gêneros: artigos de opinião; piada (monólogos); entrevistas; avisos/informativos e convite.

A partir das diversas fontes e indicadores de gêneros estabelecemos uma matriz de dados que pudesse indicar os principais gêneros a serem buscados na composição do corpus piloto da pesquisa.

Padrões utilizados em corpora de línguas faladas com escrita	Corpus português	acadêmico: enciclopédia e artigos científicos ficção/ literatura: livros de drama, suspense, romance , terror etc. oral: notícias de jornal/ entrevistas notícias: notícias de jornal
	Corpus ac/dc	artigos científicos , teses e dissertações, anais de congresso, roteiros, diversos, wikipédia, narração de jogos de futebol, manuais, revistas, jornais , horóscopo, entrevistas , diversos, contos , crônicas , variados, biografias, bulas de remédio
Padrões utilizados em corpus de línguas de sinais (mais recorrentes)	conversação/diálogos/debates narrativas (histórias pessoais, contos, crônicas, discurso sobre um tema específico, piadas) entrevistas	

Padrão de metadados IMDI	resenha de livro, gramática , romance , drama , poesia , canções, oração, catecismo, ritual de cura, ensaio político, relatório científico, discussão política , documentário, teatro, anotações pessoais
Levantamento feito com militantes da comunidade surda	narrativas surdas , contos de histórias infantis, piadas surdas, Literatura surda (poesia), Bate papo entre surdos
Levantamento de teses e dissertações	conversação natural , narrativas , literatura (piadas , contos , histórias infantis), Editais
Levantamento na plataforma de pesquisa	artigos de opinião , piada (monólogos) , entrevistas avisos/ informativos , convite

3.5 Definição dos gêneros e montagem do corpus

Nesta seção apresentaremos a parte final da aplicação e discussão de procedimentos metodológicos: a compilação do corpus. Iniciaremos com uma breve descrição tipológica do gênero, em seguida a descrição e análise dos procedimentos de levantamento de amostras do gênero e embora não esteja situado no cerne do objetivo da pesquisa, faremos algumas considerações sobre a análise tipológica e funcional dos vídeos.

Baseado na recorrência apontada pela matriz de dados anterior, definimos sete principais gêneros a serem compilados na amostragem retirada do corpus youtube: piadas, poesia, entrevistas, avisos/informativos, artigos de opinião, histórias pessoais e conversação.

Para a seleção da busca dos gêneros na plataforma utilizou-se os seguintes filtros:

- busca pelos verbetes (gênero textual) + libras + surdos

Após a amostragem do primeiro resultado para cada gênero, selecionamos os cem primeiros vídeos para análise. Retiramos da análise vídeos que não se enquadravam no tema sugerido e reduzimos a amostra para usuários nativos da língua. Uma análise qualitativa do vídeo permitiu verificar se o usuário era surdo ou um falante de segunda língua, conforme critérios já definidos na seção em que se discutiu o levantamento inicial para a busca libras. O critério “nativo” da língua, no entanto, não é tão fácil de definir, para a situação de falantes de língua de sinais. Conforme apontam Costello, Fernández e Landa (2006, p.340), as línguas de sinais fazem parte de “um contexto lingüístico não-padrão” e “têm um perfil sociolingüístico complexo e uma população de usuários totalmente atípica e heterogênea”. A primeira consideração é o fato de que os surdos são, em sua maioria, filhos de pais ouvintes, e o número de 90 a 95% tem sido apontado como um número padrão para a estatística dessa população. Esse número, no entanto, reflete a realidade norte-americana, que pode variar bastante em razão da população estudada conforme apontam os autores. O fato é que a grande maioria dos pesquisadores tem utilizado esses dados em referência a suas pesquisas com a língua de sinais do seu país para subsidiar algo que a observação participante tem atestado nas diversas comunidades surdas ao redor do mundo: o número de surdos filhos de pais surdos é ínfimo em relação à população de surdos de uma localidade. Casos específicos de populações em que por uma questão hereditária ou consanguínea o número de surdos é acima da média, como a de Marthas Vineard³⁷ nos EUA ou a Urubus Kaapor³⁸ no Brasil, se excluem dessa regra.

³⁷ Ilha localizada na costa de Massachusetts, EUA, em que por um alto índice de surdez hereditária, no início do séc XVIII até por volta de 1952, quase todos os habitantes locais conheciam e utilizavam a língua de sinais ali. Em 1854, quando a taxa de surdez na ilha atingiu o seu pico, a taxa nacional de surdez nos EUA era de 1 para 5728, enquanto que em Martha's Vineyard era de 1 para 155. Em algumas localidades da ilha, quase um quarto da população era de surdos. Um estudo sobre a ilha pode ser obtido no livro de Nora Ellen Groce, *Everyone Here Spoke Sign Language*.

³⁸ Os caapores, também chamados urubus, urubus-caapores, cambôs e ka'apor, são um povo indígena que vive no estado do Maranhão, no Brasil e que por ter uma

Considerando um número tão pequeno de indivíduos nesse status linguístico, há de se questionar, portanto, se a língua utilizada por esse grupo pode ser considerada amostra representativa da língua ou apenas de uma variação dessa. Partindo ainda de um tipo de corpus em que não se busca uma população específica, como o corpus de gêneros aqui compilado, podemos estabelecer um critério de natividade bem menos restritivo. Adotamos como critério de natividade para o corpus de gêneros-web um dos critérios elencados por Garauv e Ratman (2006): surdos, que tenham tido o contato diário com uma língua de sinais na comunidade surda por mais de dez anos, considerando aqui no entanto um tempo de 5 anos como tempo mínimo de contato.

A partir da definição dos gêneros criou-se em 2011 um canal permanente na plataforma youtube em que o corpus pudesse ser organizado e constantemente atualizado. Está organizado primeiramente em listas de reprodução dos arquivos originais. Os vídeos tratados e autorizados são repostados em listas com acesso restrito mediante solicitação individual pelo próprio canal ou pelo email corpusdelinguadesinais@gmail.com. O canal pode ser acessado no endereço <http://www.youtube.com/corpuslinguadesinais>. Os vídeos tratados e selecionados nessa pesquisa constituem um corpus web de 126 vídeos, divididos em 6 grupos denominados gêneros a partir da definição anteriormente explicitada que considera como principal os aspectos sociocomunicativos e funcionais do texto além da sua relatividade estrutural. Os grupos ficaram assim divididos: piada-18 vídeos; poesia-13 vídeos; entrevista-27 vídeos; avisos-24 vídeos; artigo de opinião-27 vídeos e histórias pessoais-17 vídeos.

A primeira busca foi para o gênero textual “piada”. Esse tipo de texto é centrado numa tipologia narrativa. Quase sempre se baseia em um monólogo narrativo de fatos fictícios, com informação de tempo e lugar fictício e envolvendo personagens caricatos do mundo real. Predominância do tempo verbal do passado, o enredo é curto, começo, meio e um final inusitado que produz o efeito humorístico. O inusitado se caracteriza em geral pela oposição daquilo que se esperava comumente para uma situação típica, oposição centrada na ambiguidade que o fenômeno linguístico proporciona comumente, expondo inclusive a problemática que a linguagem pode trazer a partir da ambiguidade, conforme define Possenti (2002). O autor pontua ainda, além do material

linguístico, a importância da análise do gênero piada por permitir observarmos temas socialmente controversos (comum ao gênero), a apresentação de estereótipos e a explicitação de discursos clandestinos ou que são comumente mascarados nos discursos oficiais. Cabe mencionar que o gênero ocupa um importante papel na composição do repertório da literatura surda brasileira, como pontua também Schallenberger (2010).

Fazendo a busca pelos termos “piada + surdo + libras” a plataforma aponta aproximadamente 14300 resultados, conforme vemos na figura a seguir:

Figura 21 – Visualização do resultado de busca do gênero piada



Fonte:

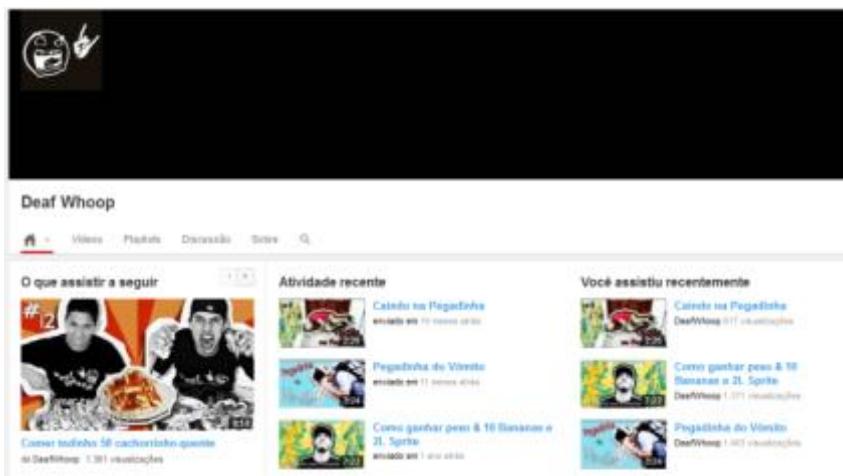
https://www.youtube.com/results?search_query=piada+%2B+surdo+%2B+libras

Analisando os cem primeiros vídeos da amostra, percebemos que a grande maioria trata-se de vídeos em que o autor filma algum evento do cotidiano, como festas, churrasco, aniversários, etc. Durante a filmagem do evento os participantes são convidados a falar algo, como apresentar-se, falar sobre o evento, ou algum comentário sobre o autor. Dessa forma o significado do uso da palavra “piada” aqui, estaria mais próximo daquilo que os falantes de português como primeira língua representam por palavras como *engraçado* ou *divertido*. Embora em alguns vídeos os indivíduos filmados façam comentários engraçados, o objetivo principal desses vídeos não é promover o humor, fazer com que o espectador do vídeo ria ou se divirta ao assistir. O objetivo desses vídeos é registrar e explicitar cenas do cotidiano do autor, seja para ser visualizado pelos participantes do vídeo, seja para que seus amigos, familiares e conhecidos possam conhecer um pouco do seu cotidiano. De fato, não se espera que desconhecidos tenham interesse em cenas cotidianas o que definiria as pessoas do círculo do autor como o público alvo, prioritariamente aqueles envolvidos no evento filmado.

A incidência principal aqui seria de vídeos que se enquadram na categoria “cotidiano”, como já descrito no levantamento inicial para a busca “surdos”. Em tese, seriam excelentes materiais para a pesquisa linguística, por apresentarem indivíduos em suas atividades diárias mais comuns. No entanto, pelo discurso não estar na centralidade do objetivo da filmagem, mas concorrer com a descrição e exposição do evento, em geral o registro dos enunciados fica extremamente prejudicado. Em grande parte deles os surdos que interagem são filmados em “recortes” das suas falas, não sendo possível recuperar os contextos anterior e posterior a fala. Considerando, portanto, que a finalidade desses vídeos não é humorística, eles foram excluídos da compilação da categoria “piada”.

O segundo grupo de vídeos mais recorrente na seleção dos cem primeiros vídeos da amostragem foram vídeos produzidos com a finalidade humorística através de esquetes teatrais. Trata-se da encenação de cenas de humor planejadas, com o uso de acessórios cenográficos, preparação de cenário, inserção de efeitos especiais e representação de atores e atrizes amadores surdos. Analisando os vídeos que apareceram no resultado e se enquadravam nesse tipo de humor, chegamos a canais “especializados”. Trata-se de canais montados por usuários surdos que buscam produzir humor para surdos em libras usando dos recursos descritos. Um deles é o canal “Deaf Whoop” que possui cerca de mil inscritos e aproximadamente cerca de 1 milhão e 200 mil visualizações. Inspirado no humor do tipo trágico-bizarro norte americano, os autores do canal apresentam vídeos em que buscam divertir o espectador permitindo observar a reação das pessoas a cenas bizarras ou a aflição dos próprios autores dos vídeos ao se submeterem a situações incômodas. Como o objetivo final é produzir humor não pela elaboração de um texto humorístico, quase não há sinalização nos vídeos, com exceção do momento inicial dos vídeos em que explicam o que será apresentado. A seguir pode-se visualizar a página principal do canal:

Fig 22 – Visualização do Canal “Deaf Whoop” no youtube



Fonte: <https://www.youtube.com/user/deafwhoop>

Tipo de vídeo	Descrição	
Humor encenado	Um ou mais surdos encenam uma piada ou situação engraçada de forma teatral com o uso de maior ou menores recursos cenográficos. Em geral são vídeos editados, com efeitos e às vezes com legendas.	19
Cenas engraçadas do cotidiano	Vídeos em que amigos e familiares do postador são filmadas em eventos cotidianos, como. Nestas filmagens alguns fazem comentários engraçados sobre alguém ou sobre a situação. Outras vezes essas pessoas estão em situação engraçada, ou eventos engraçados são filmados, como o agir de um animal de estimação ou um pequeno acidente doméstico.	55
Piadas tradicionais	Estilo tradicional de piada em que um narrador apresenta em um monólogo uma	11

	narrativa com um final inusitado	
Imitações	Esquetes em que um surdo imita a expressão corporal de outros surdos	1
Não se enquadram no tema	Vídeos que não tem a finalidade humorística, tratam-se de surdos ou sinalizadores abordando outros temas em língua de sinais	8
Erros	Vídeos que não são produzidos em língua de sinais	6

Do gênero piada então, optamos pela escolha do gênero tradicional, em que um orador apresenta em um monólogo um texto em que há um final inusitado que produz o efeito humorístico. As informações de metadados incluídas para cada vídeo compilado, tanto no gênero piada como nos gêneros seguintes foram: sexo, idade no vídeo, idade e comunidade surda com a qual aprendeu língua de sinais, comunidade surda com a qual convive atualmente, nível de escolaridade, que língua(s) utiliza para comunicação com ouvintes que não sabem língua de sinais, se tem pais ou algum parente surdo que tem ou teve convívio regular e interação através de língua de sinais, tipo de suportes em que é espectador de língua de sinais (programas interpretados, softwares de comunicação por vídeo, associação de surdos, aquisição de vídeos e materiais), frequência a algum grupo específico de surdos (associação de mulheres surdas, grupos religiosos, negros surdos, surdos gays, grupos de esportes específicos). Além disso confirmou-se com o autor a finalidade do vídeo, o público para o qual era destinado, o contexto da produção do vídeo e o objetivo final do texto. 61 vídeos se enquadraram na categoria de piadas narradas. Foram excluídos da amostra vídeos em que o autor não se enquadrava no critério de natividade, não era possível a visualização clara dos sinais, vídeos produzidos pelo mesmo autor de outro vídeo já selecionado e vídeos em que o autor era menor de dezoito anos. Dezoito vídeos foram compilados no corpus. A seguir apresentamos o endereço para a visualização desses vídeos:

- 1 - <http://www.youtube.com/watch?v=QM372k4tBIQ>
- 2- http://www.youtube.com/watch?v=xld9vj67_lo

- 3- <http://www.youtube.com/watch?v=KTJtscBoNnc>
- 4- <http://www.youtube.com/watch?v=8FO9rHERRg8>
- 5- <http://www.youtube.com/watch?v=697SwM6U-9I>
- 6- <http://www.youtube.com/watch?v=u0fTLAmCbnY>
- 7- <http://www.youtube.com/watch?v=8tDPAQ6V2qk>
- 8- <http://www.youtube.com/watch?v=np1uf3CZFiI>
- 9- <http://www.youtube.com/watch?v=nU6wburD7lg>
- 10- <http://www.youtube.com/watch?v=vw-fzcuerto>
- 11- <http://www.youtube.com/watch?v=1wcc5d3pgty>
- 12- <http://www.youtube.com/watch?v=kaagczvux4>
- 13- <http://www.youtube.com/watch?v=k9endmhtzyq>
- 14- <http://www.youtube.com/watch?v=wzf>
- 15- <http://www.youtube.com/watch?v=4gcd90jhkj0>
- 16- <http://www.youtube.com/watch?v=U1VeWiVwf-Y>
- 17- <http://www.youtube.com/watch?v=Apje5BQMKg0>
- 18- <http://www.youtube.com/watch?v=GCCcvOhzjCA>

Destaca-se o fato de 80% dos vídeos compilados serem compostos por informantes de nível de escolaridade superior e quase 90% do sexo masculino. Aproximadamente 20% dos informantes possuíam parentes surdos com o qual interagem em língua de sinais. Quinze dos dezoito vídeos trazem a temática das relações ouvintes e surdos no discurso de fundo da piada e todos eles apresentam as características tipológicas esperadas para o gênero com dois deles buscando efeito humorístico mais na performance do que no desfecho da piada.

A busca pelo gênero poesia, centrava-se no uso estético da linguagem, de maneira a produzir emoção. Essa categoria mais ampla poderia incluir poemas (estruturado em versos e/ou estrofes, rima e métrica). Observa-se uma tipologia mais narrativa ou descritiva no gênero. Essas marcas poéticas na língua de sinais brasileira podem ser percebidas principalmente, segundo Machado (2013), pelas metáforas simbólicas que produzem uma quebra na regularidade estética dos sinais e pelo que a autora chama de neologismos poéticos, sinais novos, desconhecidos pelos usuários surdos da língua e que representam esteticamente um novo conceito.

A busca inicial pelo gênero poesia foi feita através da inclusão dos verbetes “poesia + surdo + libras”. A análise dos cem primeiros vídeos possibilitou a sistematização dos resultados apresentados no quadro abaixo:

Tipo de vídeo	Descrição	
Ouvintes produzindo poesia	Vídeos em que ouvintes aprendizes ou intérpretes apresentam traduções de poesias escritas em português para língua de sinais	23
Surdos produzindo poesia	Vídeos em que surdos apresentam poesias em língua de sinais	22
Surdos traduzindo poesia	Vídeos em que surdos fazem adaptação ou tradução de poesias da língua portuguesa para língua de sinais	16
Surdos e ouvintes sinalizando outros gêneros	Surdos e ouvintes fazendo apenas referência à poesia ou sinalizando em outros gêneros	28
Surdos e ouvintes em produção técnica sobre poesia	Surdos e ouvintes realizando trabalhos técnicos sobre poesia em que há menções ou execuções de excertos de poesia apenas para análise técnica	11

O alto índice de poesias traduzidas ou adaptadas da língua portuguesa (44 vídeos), quase 50% da amostra, parece indicar uma escassez de trabalhos originais em língua de sinais. O que também é observado no alto índice de ouvintes produzindo ou traduzindo poesias em libras. Dos 22 vídeos encontrados de surdos produzindo poesias, através da análise dos canais foi possível encontrar alguns canais específicos, como o canal do poeta surdo Alan Henry, que embora use seu canal também para postagens políticas e cotidianas, publica regularmente poesias em libras, já tendo um total de oito vídeos no canal. Abaixo podemos visualizar a página de entrada do canal:

Fig 23 – Visualização do canal do poeta Alan Henry



Fonte: <https://www.youtube.com/user/alahenry>

Outra conhecida poetisa surda, Rosani Suzin, possui um canal gerenciado pelo seu esposo em que quatro de suas poesias estão publicadas. Curiosamente, o famoso poeta surdo Nelson Pimenta teve apenas uma de suas poesias no resultado da pesquisa e através da análise do seu canal, vimos que o mesmo possui apenas dois vídeos de poesia publicados. O mesmo aconteceu com a poetisa surda Fernanda Machado, que embora possua em seu canal uma playlist intitulada “antologia poética” em que reúne vinte poesias em língua de sinais de outros autores, não possui nenhum vídeo de poesia de sua autoria no seu canal, o que pode indicar uma inconsistência na realidade de produção de literatura em relação àquela retratada no suporte.

Incluindo a totalidade de poesias produzidas por surdos encontradas e aquelas disponíveis no mesmo canal dos autores, além de outras quatro poesias encontradas em canais relacionados, levantamos um total de quarenta e sete vídeos de poesias. Retirando aquelas em que o autor não se enquadrava no critério de natividade, não era possível a visualização clara dos sinais, vídeos produzidos pelo mesmo autor de outro vídeo já selecionado e vídeos em que o autor era menor de dezoito anos, chegamos a um total de treze vídeos compilados, conforme visualizamos nos endereços a seguir:

- 1-<http://www.youtube.com/watch?v=b8p8v0bvjlg>
- 2-<http://www.youtube.com/watch?v=5aegl-9bvdu>
- 3-http://www.youtube.com/watch?v=ejzme-_8cr4
- 4-http://www.youtube.com/watch?v=j_qctohicbi
- 5-<http://www.youtube.com/watch?v=5f6i6q2i8p4>
- 6-<http://www.youtube.com/watch?v=k399dqf9xri>
- 7-http://www.youtube.com/watch?v=t9ss_fcdvqa
- 8-<http://www.youtube.com/watch?v=bgrhmdbqjis8>

- 9-<http://www.youtube.com/watch?v=pxte9ahsjd8>
- 10-<http://www.youtube.com/watch?v=168sqaqljmc>
- 11-<https://www.youtube.com/watch?v=g4en44vx5qe>
- 12-https://www.youtube.com/watch?v=cqs7k_4enge
- 13-<https://www.youtube.com/watch?v=3n2xbu8nmt0>

Cabe destacar que 100% dos informantes do gênero possuem nível superior e apenas duas são informantes do sexo feminino. A recorrência nos estados representados pelos informantes também merece ser salientada. Cinco informantes são da comunidade surda de São Paulo, dois de Santa Catarina e dois do Rio de Janeiro. Em doze vídeos a presença de quebra na regularidade estética dos sinais proposta por Machado (Ibid) através da alteração na configuração de mão pode ser observada.

O gênero entrevista é marcado tipologicamente por marcas da oralidade. É classificado como do tipo dialogado e em alguns corpora é inserido no gênero conversa. Elementos da interação linguística compõe o texto como a interlocução, perguntas e respostas, fala diretamente determinada pelo enunciado do interlocutor, frases breves e às vezes sobrepostas, elipses, onomatopéias, interjeições, e elementos dêiticos. Em geral se constroem numa estrutura de cumprimentos, preparação, desenvolvimento e despedida. Apresenta também características de texto dissertativo expositivo, pela exposição de ideias objetivas na busca de informar algo. Também pode apresentar características argumentativas quando se entrevista alguém com o objetivo de saber sua opinião sobre algum tema específico. O objetivo é através de uma dissertação expositiva informar o inquiridor e ao público prioritário que irá assistir, quem é o entrevistado e o que ele pensa em relação a algo. Por vezes há alternância entre uma narrativa detalhada, quando se pede que o entrevistado fale sobre sua trajetória de vida até uma argumentação mais elaborada, quando se pede que ele se posicione em relação a um tema.

Para a busca desse gênero então, inserimos os verbetes “entrevista + surdo + libras”. Os resultados foram dez vídeos com surdos sendo entrevistados. No entanto somente duas entrevistas eram realizadas por outros surdos. Seis foram realizadas por ouvintes aprendizes de libras e duas por ouvintes não sinalizantes com auxílio de intérpretes. Foram desconsideradas duas entrevistas realizadas por surdos em outras

línguas de sinais (língua gestual portuguesa e sinais internacionais). Os outros noventa vídeos foram classificados da seguinte maneira, de acordo com o conteúdo:

Tipo de vídeo	Descrição	
Surdos e ouvintes sinalizando outro gênero	Vídeos em que surdos e ouvintes sinalizam em gêneros como narrativas, piadas, jornal e outros que não se enquadram no gênero entrevista	37
Ouvintes dando entrevista sobre surdos/surdez	Ouvintes dando entrevista em língua portuguesa sobre surdos ou surdez. 12 entrevistas tinham interpretação em libras, 9 eram apenas em língua portuguesa e em 6 os ouvintes sinalizavam	27
Ouvintes simulando entrevista em libras	Ouvintes simulando uma atuação em língua de sinais no gênero entrevista para a realização de trabalhos de curso de libras	14
Surdos dando entrevista em língua portuguesa	Entrevistas originalmente veiculadas na televisão em que famosos surdos sinalizantes e oralizados da comunidade davam entrevista em língua portuguesa através da leitura labial e oralização.	4
Erros	Resultados que nada tinham a ver com libras	4
Surdos dando entrevista em outras línguas de sinais	Surdos sendo entrevistados em língua gestual portuguesa e em sinais internacionais	2
Orientações sobre entrevista para surdos	Vídeos em que são feitas orientações de entrevista de emprego para surdos, um em libras e outro com o uso de desenho animado	2

Como o número da amostra foi considerado pequeno, a partir dos resultados da análise, buscou-se novos vídeos no gênero entrevista, primeiro por um levantamento nos canais dos vídeos já encontrados e depois por levantamento em canais que estavam relacionados a esses, semelhante ao que foi feito no gênero piada. O primeiro resultado da busca, por exemplo, trata de uma entrevista de um professor surdo postada no canal vlogletraslibras³⁹. Fazendo uma busca no próprio canal através da palavra “entrevista” encontramos outra entrevista realizada com uma professora surda, conforme se vê na figura a seguir:

Fig 24 - Visualização da busca por “entrevista” no canal vlogletraslibras



Fonte: <https://www.youtube.com/user/vlogletraslibras?feature=watch>

Utilizando essa nova estratégia de busca foi possível compilar mais trinta vídeos em que surdos eram entrevistados. Seis vídeos em que os surdos apenas respondiam mas não era possível ver as perguntas em língua de sinais foram retirados da análise. Outros quatro vídeos com surdos dando entrevista oralizando também foram excluídos. Por último, três vídeos que tratavam de filmagens de entrevista de surdos em outros suportes (filmagem da projeção da entrevista no telão, filmagem da TV) foram retirados por não permitirem visualizar os dados com nitidez. 17 novos vídeos foram incluídos no gênero, totalizando 27 vídeos de entrevistas analisados, conforme visualiza-se a seguir:

³⁹ Disponível em <http://www.youtube.com/user/vlogletraslibras>

- 1-<http://www.youtube.com/watch?v=bp3WOdNzON0>
- 2-<http://www.youtube.com/watch?v=WJMtiommb3w>
- 3-<http://www.youtube.com/watch?v=fBIzjjmAjA0>
- 4-<http://www.youtube.com/watch?v=l9XTuZVXhCY>
- 5-http://www.youtube.com/watch?v=_H_hlx2C-KI
- 6-<http://www.youtube.com/watch?v=EsNGxhbhTWE>
- 7-<http://www.youtube.com/watch?v=k9duvnpw5aQ>
- 8-<http://www.youtube.com/watch?v=p3QPSQEzGMA>
- 9-<http://www.youtube.com/watch?v=elkckbaa8qq>
- 10-<http://www.youtube.com/watch?v=V3L7P2o5rq0>
- 11-<http://www.youtube.com/watch?v=ZDNuCu3Pc3o>
- 12-<http://www.youtube.com/watch?v=6eotVw9EBkE>
- 13-<http://www.youtube.com/watch?v=Eki6KXhhEGw>
- 14-<http://www.youtube.com/watch?v=XxPsXUF0XIM>
- 15-<http://www.youtube.com/watch?v=ZBRDaqNRi8>
- 16-<http://www.youtube.com/watch?v=y-zojhj5q5s>
- 17-<http://www.youtube.com/watch?v=uC9ZZVcmh9o>
- 18-http://www.youtube.com/watch?v=2hCDwS_fkyc
- 19-<https://www.youtube.com/watch?v=B9R2P9Pjo9c>
- 20-https://www.youtube.com/watch?v=gRkME_Okl34
- 21-<https://www.youtube.com/watch?v=oacU4hot9Z0>
- 22-https://www.youtube.com/watch?v=4wevio__TwE
- 23-https://www.youtube.com/watch?v=_w9LGsPO6ys
- 24-<https://www.youtube.com/watch?v=j85hJZOBmtA>
- 25-<https://www.youtube.com/watch?v=WCPLKKWEdsM>
- 26-<https://www.youtube.com/watch?v=KpcTxmmFoYE>
- 27-<http://www.youtube.com/watch?v=IZEKjPfmtqc>

Aproximadamente 30% dos vídeos são entrevistas realizadas por aprendizes de libras. Destaca-se também a diversidade de estados obtidos na amostra: onze. Predominou-se a tipologia dissertativa-expositiva, pela exposição de ideias objetivas na busca de informar algo.

Uma configuração específica de comunicação, transmitir uma informação direta, geralmente breve e objetiva, para um grupo específico ou para todos os espectadores surdos do vídeo, configuram o gênero textual aviso. Tipologicamente expositivo objetivo, podendo ser enquadrado também numa tipologia dissertativo-expositivo, quando além de informar o sinalizador inclui julgamento de valores sobre o evento. Pode incluir características descritivas ao detalhar um endereço ou mapa por exemplo. Se o texto incluir detalhamento de como realizar uma ação também pode apresentar características injuntivas, como num aviso de

recomendações de trânsito. A busca foi feita pelas palavras “aviso + surdo + libras”.

O levantamento apontou aproximadamente 7410 vídeos compatíveis com a busca. Os resultados foram os mais variados. Dos 100 primeiros vídeos analisados, encontramos os seguintes dados:

Tipo de vídeo	Descrição	
Surdos sinalizando em artigos de opinião	Vídeos em que surdos sinalizavam sua opinião sobre algum tema específico. A intenção aqui era do uso do sinal de aviso no sentido de divulgar, publicar sua opinião. Os vídeos foram enquadrados no gênero artigo de opinião	14
Surdos sinalizando avisos diretos	Vídeos que continham avisos diretos. Podiam ser divididos em dois grupos. O primeiro, avisos curtos de 1 a 3 minutos, como orientações sobre a roupa para se usar numa festa, apresentação da nova diretoria da associação, informação sobre um canal de contato de uma instituição e outros do tipo. Também poderiam ser destinados aos espectadores surdos em geral, como recomendações de segurança na internet, ou informações sobre como usar o passe livre nacional. O segundo, avisos mais detalhados, média de 3 a 5 minutos, como orientações para candidatas surdas de um estado sobre como proceder na inscrição de um concurso, orientações para os surdos de outro estado sobre como tirar a carteira de motorista, e outros do mesmo tipo com alguns traços injuntivos.	12
Surdos sinalizando tutoriais	Vídeos que se enquadravam na categoria injuntivo, consistindo de orientações diretas e didáticas, passo a passo, de como proceder em uma situação específica.	4

	Esse tipo de vídeo foi enquadrado no gênero tutorial	
Surdos sinalizando convites	Vídeos que apresentavam avisos do tipo convite, em que informações sobre um evento são dadas (dia, hora, objetivo) e os espectadores são convidados a participar.	14
Surdos sinalizando outros gêneros	Vídeos que apresentavam surdos em outro gênero: piada, palestra, falando sobre seu cotidiano, apresentando uma cena cotidiana, ensinando sobre cultura surda, dando entrevista, contando uma piada, ou estória infantil	44
Ouvintes sinalizando	Vídeos que apresentavam ouvintes convidando ou orientando surdos sobre algo específico	8
Erros	Eram resultados errados, que nada tinha a ver com libras	4

Consideramos, então para a análise, avisos diretos, do tipo curto ou longo, destinado a público específico. Entrando nos canais apresentados nos resultados iniciais, encontramos outros vinte vídeos semelhantes nos resultados e canais relacionados. Retirando aqueles que não se enquadravam nos padrões de compilação já mencionados nos gêneros anteriores, restaram doze novos vídeos, totalizando vinte e quatro vídeos compilados.

Para o gênero avisos então, tivemos como resultado:

- 1-http://www.youtube.com/watch?v=WeC4yF-NJdY&list=PLgR-wnZGYwFd8JRbYRoAo9mo6H3LSZY_J
- 2-<http://www.youtube.com/watch?v=NucF-uwyeEA>
- 3-http://www.youtube.com/watch?v=zU4He_yUbY4
- 4-<http://www.youtube.com/watch?v=WpT4ZNzmgmU>

5-<http://www.youtube.com/watch?v=d8UFyz4qDvI>
6-<http://www.youtube.com/watch?v=UJGyMXZmO4c>
7-<http://www.youtube.com/watch?v=ygu3uFNpr5k>
8-http://www.youtube.com/watch?v=Ll6fG_aTDqU
9-<http://www.youtube.com/watch?v=z9tzNIL7GmA>
10-<http://www.youtube.com/watch?v=0rQT-0KNW80>
11-<http://www.youtube.com/watch?v=eWd3iOMsQLs>
12-<http://www.youtube.com/watch?v=CSH87FJ1mps>
13-<http://www.youtube.com/watch?v=aTsWTpksM8E>
14-<http://www.youtube.com/watch?v=VaVsY2gZIEY>
15-http://www.youtube.com/watch?v=G_XID8ItjMg
16-<http://www.youtube.com/watch?v=Xn4PotudyA8>
17-<http://www.youtube.com/watch?v=3jD0Dzik59M>
18-<http://www.youtube.com/watch?v=Xa4ZLVInoZ0>
19-<http://www.youtube.com/watch?v=QMix9Df0RkE>
20-<http://www.youtube.com/watch?v=N1NqKLS1XOI>
21-<http://www.youtube.com/watch?v=YIj11fSf92U>
22-http://www.youtube.com/watch?v=AYJ_hJGNSfM
23-<http://www.youtube.com/watch?v=DWOgwbwgd-c>
24-<http://www.youtube.com/watch?v=fRT6smzP3Bg>

Aqui tivemos um índice maior, de quase 40% de informantes do sexo feminino. O nível de escolaridade superior aqui caiu para 70%. Houve uma variação maior na tipologia dos textos indo de um extremo expositivo-descritivo para textos com quase metade das sequências tipológicas dissertativo-argumentativo.

O gênero artigo de opinião consiste na defesa de idéias ou do ponto de vista do autor do texto. O autor tenta persuadir o interlocutor quanto a algum ponto de vista, utilizando-se de suposições lógicas, fatos, exemplos pessoais, confronto de idéias, testemunhos etc. Predomina-se o uso denotativo da linguagem. Uso frequente de conectores explicativos, de causa e consequência e coordenadores. Em geral, estruturam-se pela exposição da tese, demonstração e conclusão. Tende a usar uma variação padrão da linguagem. Apresenta uma função apelativa da linguagem com o objetivo de convencer. Bons artigos de opinião baseiam-se em idéias organizadas de forma clara, precisa e ordenada. Observa-se alta incidência de verbos que caracterizam opinião como: acho, acredito, vejo, sinto, percebo.

Aproximadamente 3640 resultados apareceram para a busca “opinião + surdo + libras”. Dos cem primeiros levantados, apenas nove correspondiam a vídeos em que surdos opinavam sobre algum tema específico. No entanto, esses vídeos levavam, pela exibição de vídeos relacionados a uma grande rede de vídeos discutindo o mesmo tema. Em geral, temas polêmicos como implante coclear, educação de surdos e relação ouvintes e surdos. A maioria dos vídeos vem de grupos de discussão do facebook em que um tema é levantado. Uma parcela dos vídeos apresenta surdos opinando sobre um neologismo ou uma variação linguística de um sinal, quase sempre discussão iniciada no grupo “dicionário libras” do facebook. Incluindo os principais vídeos relacionados, compilamos vinte e sete vídeos que se enquadram no gênero e nos critérios do corpus. Os resultados podem ser visualizados a seguir:

- 1-http://www.youtube.com/watch?v=bVhTxxGQX_8
- 2-http://www.youtube.com/watch?v=9_A17BVqbB0
- 3-<http://www.youtube.com/watch?v=ujgt-lHUt0s>
- 4-http://www.youtube.com/watch?v=Ry_xhAQR6Z4
- 5-http://www.youtube.com/watch?v=VLmB_Lg_WnY
- 6-<http://www.youtube.com/watch?v=tDsxDZ6gy9A>

Discutindo educação inclusiva e educação bilíngue para surdos

- 7-<http://www.youtube.com/watch?v=bvca7vubdZ0>
- 8-<http://www.youtube.com/watch?v=8VifwuwE3CQ>
- 9-<http://www.youtube.com/watch?v=48Jumb59BZA>
- 10-<http://www.youtube.com/watch?v=eVmgdFWK3jg>

Discussão sobre implante coclear

- 11-<http://www.youtube.com/watch?v=pTvZFJaeoIQ>
- 12-<http://www.youtube.com/watch?v=vjWmEkAspz4>
- 13-<http://www.youtube.com/watch?v=YcoNRNO8nbQ>
- 14-<http://www.youtube.com/watch?v=ZxiOPCLH8oI>
- 15-<http://www.youtube.com/watch?v=jb1DRyNDaTE>
- 16-<http://www.youtube.com/watch?v=8f9FnlZxhdE>
- 17-http://www.youtube.com/watch?v=l0h4_BIIyDY
- 18-<http://www.youtube.com/watch?v=cQXiwk4MK7A>
- 19-<http://www.youtube.com/watch?v=outrnhkmebo>

Discussão sobre o uso do sinal “I-LOVE-YOU”

- 20-<http://www.youtube.com/watch?v=e44w0po-IBE>
- 21-<http://www.youtube.com/watch?v=iN1L6zpcJvQ>
- 22-<http://www.youtube.com/watch?v=82TuT9oovWk>

23-<http://www.youtube.com/watch?v=k7kSz4OFizU>
 24-<http://www.youtube.com/watch?v=cfjIaooAyWc>
 25-<http://www.youtube.com/watch?v=k36nzCVUo5k>
 26-<http://www.youtube.com/watch?v=n0L8jF0Qwno>
 27-<http://www.youtube.com/watch?v=BkSDtjxW9Co>

As histórias pessoais aparecem geralmente em categorias como “narrativas” nos corpora de línguas. De tipologia narrativa, em geral os vídeos que classificamos nesse grupo apresentam a história cronológica de vida do autor de forma resumida até o momento em que a narração é feita ou descrevem um recorte específico de tempo de sua vida ou de um acontecimento que presenciou/participou. Em geral utilizam-se de verbos no passado, verbos de ação, conectores cronológicos substantivos, adjetivos e advérbios de lugar. Estruturam-se em: apresentação, núcleo e parte final. Quase sempre se usa uma variação padrão da língua a partir de uma função referencial.

A busca por “história + surdo + libras” trouxe apenas cinco resultados em que surdos narravam a história de suas vidas ou momentos específicos de sua história. Foi necessário fazer uma busca por canais pessoais relacionados aos resultados. Os outros noventa e cinco vídeos apresentaram os seguintes dados:

Tipo de vídeo	Descrição	
Ouvintes e surdos falando sobre história dos surdos	Vídeos sobre a história dos surdos ou algum aspecto da história dos surdos em formato de documentário, narrativa, montagem ou teatro. A histórias eram contadas por ouvintes ou por surdos. A grande maioria narra atividades de disciplina de graduação	57
Ouvintes e surdos sinalizando estórias infantis	Vídeos que eram estórias infantis contadas por surdos ou ouvintes	18
Ouvintes interpretando estórias	Vídeos que eram interpretações de estórias infantis	2

infantis		
Ouvintes e surdos no gênero cinema amador	Vídeos em que surdos ou ouvintes participam de um filme amador	6
Surdos outros gêneros	Vídeos que tratavam de surdos sinalizando em outros gêneros	8
Erros	Vídeos que não correspondiam à busca	4

Através da busca por canais pessoais de indivíduos surdos apresentados na amostra, encontramos outros vinte vídeos em que surdos narram aspectos da sua vida. Desses, quatro eram surdos que ingressaram a pouco há tempo na comunidade surda. Quatro deles apresentavam problemas de visualização do vídeo, por iluminação e por velocidade da captura das imagens. Os demais se enquadravam nos critérios de compilação da pesquisa, totalizando dezessete vídeos compilados:

- 1-<http://www.youtube.com/watch?v=c1g9lnprjqk>
- 2-<http://www.youtube.com/watch?v=5rqvdxgv4uy>
- 3-<http://www.youtube.com/watch?v=j5q1senorlo>
- 4-<http://www.youtube.com/watch?v=jiaojooykss>
- 5-http://www.youtube.com/watch?v=tgst6_9zkxs
- 6-<http://www.youtube.com/watch?v=y9wbyg2qzue>
- 7-<http://www.youtube.com/watch?v=3mlkl7emxcs>
- 8-<http://www.youtube.com/watch?v=5tzknmq1gim>
- 9-<http://www.youtube.com/watch?v=lzivk4x5jxy>
- 10-<http://www.youtube.com/watch?v=whbacsz-6yi>
- 11-<http://www.youtube.com/watch?v=weq1fdds5lu>
- 12-<http://www.youtube.com/watch?v=kv-rivcr2xy>
- 13-<http://www.youtube.com/watch?v=nqxmbzbe5py>
- 14-http://www.youtube.com/watch?v=eaec-9gwx_u
- 15-<http://www.youtube.com/watch?v=jbto30skfcq>
- 16-<http://www.youtube.com/watch?v=h8f10g9f0sm>
- 17-<http://www.youtube.com/watch?v=t5qx3c6uxt8>

De tipologia dialogal, o gênero conversa apresenta características próprias da oralidade. A variação coloquial é predominante. O tema em discussão é construído pela interação entre os interlocutores. Os objetivos e estratégias de comunicação são reconstruídos a todo momento a partir da observação da compreensão do interlocutor. Por vezes há sobreposição de falas, interrupções, abandono de idéias e realocações de temas e propostas discursivas. A busca esperava encontrar vídeos em que surdos estivessem interagindo em conversas informais em que material linguístico significativo pudesse ser analisado.

A busca por “conversa + surdo + libras” trouxe aproximadamente 7420 resultados. Dos 100 primeiros vídeos da amostra, apenas seis deles apresentavam surdos interagindo em conversas espontâneas. Quatro deles não tinham qualidade suficiente de imagem para que a conversa pudesse ser analisada. Um grupo significativo de resultados no entanto, vinte e seis, apresentavam surdo interagindo em programas de conversa por webcam, como skype, oovoo, camfrog, dentre outros. A qualidade da sinalização nessas gravações, alto índice de adolescentes nas imagens (quatorze vídeos) e usuários que não se enquadravam no critério de nativo do corpus fez com que a compilação desse gênero fosse descartada.

As estratégias para a compilação do corpus-web através do gênero apresentaram algumas particularidades. Na decisão de busca pela nomenclatura comum do gênero juntamente com o status do falante - surdos, e a nomenclatura da língua utilizada - libras, uma grande quantidade de vídeos que não se enquadravam no tema retornaram nos resultados. Em média, apenas 10% dos vídeos da amostra se enquadravam na busca. Uma busca que se restringisse aos vídeos que continham os três nomes no título, do tipo “allintitle”, não apresentou resultados satisfatórios. Para a busca do gênero poesia em libras, por exemplo, apenas dois supostos vídeos apareceram, conforme vemos na figura a seguir:

Fig 25 - Resultado da busca “allintitle” para o gênero poesia em libras



Fonte:

https://www.youtube.com/results?search_query=allintitle+poesia+surdo+libras

Após a análise dos vídeos acima vimos que nenhum dos dois se enquadravam no gênero, sendo o primeiro um aviso sobre uma oficina e o segundo um trabalho feito por aprendizes. Se considerássemos somente a busca por *poesia + libras* teríamos quatorze resultados, onde seis tratavam-se de poesias feitas por surdos. Na busca por *allintitle poesia surdos*, os resultados apresentaram seis vídeos, sendo que quatro vídeos referiam-se a poesias de surdos. As duas buscas restritas ainda foram muito inferiores aos resultados encontrados em uma busca por canais relacionados e redes de relacionamentos entre canais. A primeira busca por canais, “*poesia libras, channel*” trouxe nos cem primeiros resultados, canais que continham vinte e duas das cinquenta poesias encontradas na busca mais detalhada feita na pesquisa, e a busca “*poesia surdo, channel*” trouxe nos cem primeiros canais, trinta e nove das cinquenta poesias encontradas numa busca completa. Os resultados da pesquisa nessa etapa sugerem uma sistematização de busca para a formação de corpus de vídeos na plataforma que coloque em primeiro plano a busca por canais, playlists e usuários, antes dos resultados por títulos. A proposta de sistematização seria: 1- busca por nomenclatura comum do gênero + canal; 2- busca pela nomenclatura comum do gênero + playlist; 3- busca no canal pela nomenclatura comum do gênero; 4- busca no canal por canais relacionados; 5-busca no canal por perfil relacionados; 6-busca nos perfis relacionados pela nomenclatura comum do gênero.

Outro aspecto importante observado foi aquele relacionado ao uso da nomenclatura comum do gênero. A escolha da palavra em portu-

guês pode representar um campo semântico mais amplo ou mais restrito que será determinado pelo conhecimento do pesquisador em relação ao uso dos sinais e suas representações em segunda língua. A busca pelo gênero *piada*, por exemplo, trouxe resultados relacionados mais a outros campos do que ao sentido literal da palavra. O mesmo aconteceu para a busca por “*conversa surdo*” ou “*história surdo*”.

Dessa forma podemos sintetizar abaixo o passo-a-passo na constituição do corpus-web:

1 - Definir o programa de extração de vídeos. Definiu-se o programa *atube catcher*;

2 - Definir a plataforma fonte de vídeos e criar uma conta para hospedar os vídeos. *youtube.com* foi a selecionada;

3 - Definir a organização das pastas do corpus, formato, local de hospedagem física e de backup dos vídeos;

4 - Na plataforma escolhida fazer a busca de acordo com a finalidade da pesquisa.

Para a busca na plataforma *youtube.com* alguns recursos podem ser utilizados como:

- Incluir aspas para forçar a plataforma a buscar exatamente o que você deseja.

ex: "piada surdo"

- Incluir *intitle*: para buscar apenas vídeos que contenham as palavras buscadas no título.

ex: *intitle*:piada surdo

- Incluir *allintitle* para que todos os termos procurados apareçam no título

ex: *allintitle*:piada surdo

- Incluir + ou - para incluir ou excluir palavras específicas nos resultados

Ex: piada surdo – ouvinte

Após a busca pelos termos inseridos, a própria plataforma oferece a possibilidade de refinar a busca clicando em “filtro”. Esse comando oferece o refinamento de busca por data do upload (última hora,

hoje, esta semana, este mês, este ano), tipo de resultado (vídeo, canal, playlist, filme, programa), duração (curto, longo), características (HD, CC, Creative commons, 3d, No ar, comprado) e classificação (por relevância, data do upload, contagem de visualizações, avaliação). O uso deste comando também pode ser feito diretamente na barra de buscas através da inserção da vírgula e do nome do comando após o termo.

Ex: piada surdo, week. Para buscar vídeos que contenham referência aos termos “piada” e “surdos” que tenham sido postados na última semana.

Piada surdo, playlist. Para buscar listas de vídeos com referência aos termos “piada” e “surdos”.

5 - Para a busca de vídeos em língua de sinais, fazer primeiramente a busca pela nomenclatura do gênero específico (piada, notícia, história etc) mais o filtro de canal. Ex: “poesia libras”, channel

6 – Em seguida fazer busca pela nomenclatura do gênero específico e o filtro de listas de reprodução. Ex: “poesia libras”, playlist

7 – Com base nas informações iniciais dos metadados inseridos pelo postador, classificar os vídeos nos resultados encontrados a partir dos recortes da pesquisa: idade, status do sinalizador, gênero, região, etc. Refinar a classificação com base em informações extras como aquelas descritas em outros vídeos no canal do postador, links externos e resultados de buscas na web relacionadas ao postador. Inserir os vídeos em listas de referência aos vídeos originais no canal da pesquisa.

8 – Com os vídeos selecionados pré-classificados, realizar contato com o postador solicitando autorização de uso na pesquisa. Após o consentimento, repostar os vídeos no canal da pesquisa.

9 – Selecionar mais vídeos afins a partir de buscas nos canais de postadores com grandes quantidades de vídeos. No canal escolhido fazer a busca por canais relacionados. No canal escolhido buscar perfis relacionados. Nos perfis relacionados buscar a nomenclatura comum do gênero

CONSIDERAÇÕES FINAIS

Algumas considerações podem ser feitas a partir da aplicação da pesquisa. Apresentamos como objetivo discutir os procedimentos metodológicos na formação de um corpus de línguas de sinais a partir da plataforma youtube, elaborando e identificando as principais questões que cercam um corpus de vídeos. Dentro dessa discussão pudemos destacar a reflexão e discussão sobre as questões éticas que trouxeram à tona questões importantes para a pesquisa em língua de sinais como a reflexão do papel do pesquisador e sua relação com a comunidade. Discussões iniciadas a pouco tempo na ética em documentação linguística puderam ser situadas no universo da pesquisa com língua de sinais, encontrando inclusive confluências com achados bibliográficos na pesquisa, como o manifesto de Amsterdam, que validam as proposições teóricas tomadas como referência. Como resultado a pesquisa indicou a necessidade de uma reelaboração dos procedimentos éticos para a pesquisa com línguas de sinais, de forma que considerem a participação de pesquisadores surdos, o uso de imagens e o impacto das mesmas e o retorno dado as comunidades envolvidas a partir das pesquisas e coleta de dados.

Questões éticas importantes destacadas por autores como Dwyer (2006), Hochgesang et al (2010), Crippen e Robinson (2013), Reppen (2005) e por outros teóricos tomados como base para a elaboração da declaração de ética da sociedade de linguística da língua de sinais (SLLS), traduzida e discutida aqui, puderam ser verificadas e aplicadas na plataforma youtube e orientações quanto a utilização ética da plataforma foram resenhadas. A discussão teórica permitiu também que fossem sistematizados procedimentos práticos para fases específicas de pesquisa, oferecendo material didático de orientação como ponto de partida para questões a serem pensadas em todo o processo. Um questionário aplicado a amostra da população acadêmica surda reforçou os encontros emergentes nas discussões teóricas entre propostas científicas e necessidades das comunidades envolvidas. A partir das diretrizes éticas também, elaborou-se um termo de consentimento de pesquisa em língua de sinais e língua portuguesa que foi postado no canal do corpus e pode ser tomado como base para a formação de futuros corpus-web.

Quanto aos procedimentos técnicos de formação do corpus, objetivávamos descobrir como coletar, analisar e referenciar um corpus de língua de sinais, verificando a aplicabilidade de procedimentos já descritos na literatura para esse trabalho. Como resultado a pesquisa apontou a possibilidade de usos de ferramentas para extração dos vídeos na internet inclusive contornando problemas comuns com arquivos de filmes desse tipo. Indicou a adequação de programas e softwares para um dos formatos mais comum de vídeos em língua de sinais na internet, o .swf. Um levantamento com várias plataformas de vídeo na internet sugeriu indicar a plataforma youtube como a mais coerente com uma pesquisa ampla de formação de corpus-web no que diz respeito a libras, seja pela quantidade de vídeos postados, pela funcionalidade, possibilidade de refinamento de busca, e indicação de pesquisas sobre o site.

A pesquisa propôs uma organização dos corpus de vídeos baseada em Sardinha (2004) e através da aplicação prática dessas orientações, demonstrou a necessidade de procedimentos específicos como: conversão dos vídeos baixados para a extensão adequada ao software escolhido para transcrição; necessidade de manutenção dos arquivos originais em seus formatos; manutenção de arquivos com extensões específicas para diretórios específicos e a necessidade do uso de um diretório externo. Problemas específicos no tratamento de vídeos, levantados por Sloetjes(2011), Bickford (2005), Leite (2008), puderam ser confirmados na organização do corpus, reforçando a orientação desses autores em cenários diferenciados para o cenário também da compilação de corpus-web.

Ao contrário da interface aparentemente caótica sugerida pela plataforma que invalidaria seu uso, a pesquisa demonstrou que estratégias podem ser desenvolvidas para a localização de metadados, autoria e retirada de informações dos vídeos. O índice de 6% de indefinição de status do autor do vídeo, no primeiro levantamento analisado, indica essa viabilidade.

Ao analisar o uso dos gêneros em corpora, verificou-se equívocos entre a tipologia e o gênero na categorização propostas por estes. Levantando os trinta e quatro projetos baseados em corpus propostos por Konrad (2011), localizou-se oito deles que poderiam ser classificados como projetos com a finalidade de constituir uma documentação linguística abrangente da língua de sinais. A interação e a tipologia narrativa mostrou ser a mais presente nos gêneros recorrentes desses projetos e dos outros suportes escolhidos para a análise. Em se tratando espe-

cificamente do levantamento na plataforma youtube a pesquisa parece indicar, tomando a amostra utilizada, que o status do falante é mais significativo para buscas de material produzido por falantes nativos do que a busca pela nomenclatura da língua.

O gênero mais recorrente tomado como indicador no levantamento utilizado, o gênero conversa e que aparentemente seria o mais utilizado no corpus de gêneros como contribuição típica do suporte web, evidenciou-se na pesquisa como inviável, já que aspectos técnicos básicos de filmagem e centralidade no discurso linguísticos não são priorizados pela grande maioria dos vídeos publicados nesse gênero. A artificialidade de alguns gêneros como o cinema amador e o humor encenado também apresentaram grau de incidência suficiente para retirá-los da análise. A relevância do gênero piada apontado por Schallenger (2010) parece se confirmar no maior índice apontado nesse gênero. Tanta as amostras levantadas nesse como nos demais gêneros demonstraram coerência com as características de gênero previstas no início de cada busca. A análise dos procedimentos metodológicos na etapa final da pesquisa indicou a necessidade de um cuidado com os termos inseridos na busca, considerando o universo pragmático diferenciado para as representações dos termos usados por ouvintes e dos usos dos mesmos termos usados por surdos como segunda língua. A busca simples por nomenclatura de gêneros mostrou-se secundária frente a resultados obtidos por busca de canais, playlist's e redes de relacionamento entre usuários e canais.

Embora não constituindo-se objetivo final da pesquisa, alguns apontamentos iniciais podem ser feitos a partir de uma análise prévia dos resultados da compilação de gêneros, como: alto nível de escolaridade dos informantes (83% de informantes de nível superior para todos os gêneros), predominância de informantes do sexo masculino (66% para todos os gêneros), filhos de surdos como constituintes de uma variação linguística mínima no corpus (6%), centralidade em amostras referentes a estados do sul e sudeste (42%). Essas e outras características a serem posteriormente analisadas, podem constituir um norte para os usos possíveis da plataforma para pesquisas específicas.

O resultado da análise exaustiva e extremamente laborosa feita nessa pesquisa, de aproximadamente 1100 vídeos constituem material primário disponível agora no maior corpus de gêneros da língua de sinais brasileira na plataforma youtube, e embora ainda necessite de um tratamento mais criterioso, possibilita que pesquisadores possam recor-

rer a material previamente classificado e autorizado para uso em pesquisa. Além das pesquisas no campo dos estudos linguísticos, estudos em outros âmbitos podem beneficiar-se dos dados que evidenciam padrões de comportamento dos usuários surdos e outros membros dessa comunidade na maior comunidade de vídeos em língua de sinais brasileira em rede. A regularidade dos padrões tipológicos e funcionais dos materiais publicados na plataforma e sua similaridade com as postagens feitas por outros grupos socioculturais parece indicar um uso não tão diferenciado da rede como se supunha inicialmente. Podemos concluir indicando a possibilidade de utilização da plataforma youtube como fonte para a formação de corpus linguístico limitada no entanto por gêneros específicos, categorias, e informantes de grupos específicos, conforme evidenciados aqui.

REFERÊNCIAS

ALMEIDA, Y. **Youtube e a revolução digital**. Resenha disponível em <<http://herdeirodoacoes.wordpress.com/2010/01/18/resumo-do-livro-youtube-e-a-revolucao-digital/>>. Acesso em 12 de dezembro de 2010.

ARMSTRONG, D. F., KARCHMER, M. A. e VAN CLEVE, J. V. **The Study of Signed Languages: Essays in Honor of William C. Stokoe**. Washington, DC: Gallaudet University Press, 2002

BAKHTIN, M. **Estética da criação verbal**. 2 ed. Martins Fontes. São Paulo. 2000

BENTES, A.C. e MUSSALIM, F. (org.). **Introdução à Linguística: fundamentos epistemológicos**. São Paulo: Cortez, 2004, vol. 3

BERBER SARDINHA, T. **Linguística de Corpus: histórico e problemática**. DELTA, São Paulo, v. 16, n. 2, 2000. Disponível em <http://www.scielo.br/scielo.php?script=sci_arttext&pid=S0102-44502000000200005&lng=en&nrm=iso.> Acesso em 10 de Julho de 2011.

_____. **Linguística de Corpus**. Barueri, SP: Manole, 2004.

BERTI, O.M.de C. **O YouTube e o fim da televisão no Brasil**, Comunicação Científica apresentada no Grupo Temático de Produção e Recepção do XIV CELACOM – Colóquio Internacional da Escola Latino-Americana de Comunicação. Memorial da América Latina, São Paulo (SP), Brasil. De 17 a 19 de maio de 2010

BIBER, D. **Representativeness in Corpus Design**. Em Practical Lexicography. A Reader. Editado por Thierry Fontenelle. Oxford University Press. New York. USA. P. 63 – 87. 2008

BICKFORD, Albert. **Using ELAN: A getting-started guide for use with sign languages**. 2005

BRASIL, **lei nº 12.853, de 14 de agosto de 2013**. Disponível em <http://www.planalto.gov.br/ccivil_03/_Ato2011-2014/2013/Lei/L12853.htm>

BURGESS, J. e GREEN, J. **Youtube e a Revolução digital**. Como o maior fenômeno da cultura participativa transformou a mídia e a sociedade SÃO PAULO, Aleph, 2009

BYBEE, Joan. 2010. *Language, usage and cognition*. Cambridge: Cambridge University Press. 252 págs

CLANCY, B. **Building a corpus to represent a variety of a language** Em The Routledge Handbook of Corpus Linguistics. O’Keeffe A. and McCarthy, M. (Ed.) Routledge. New York, NY. 2010

COSTELLO, B. FERNÁNDEZ, J. e LANDA, A. **O sinalizante nativo não-(existente): pesquisa em língua de sinais em uma pequena população surda**. 340-355. Em questões teóricas das pesquisas em línguas de sinais. 9º congresso internacional de pesquisas em línguas de sinais. Florianópolis, Brasil, 2006

CRASBORN, O. **What Does “Informed Consent” Mean in the Internet Age? Publishing Sign Language Corpora as Open Content**. Sign Language Studies, 10-2, 2010, p. 276-290. Disponível em

<http://www.ru.nl/publish/pages/515325/crasborn2010_sls_informecedconsent.pdf >

CRIPPEN, J. A. and ROBINSON, L.C. 2013. **In Defense of the Lone Wolf**: Collaboration in Language Documentation. Language Documentation & Conservation. 7:123-135. Disponível para download em

<<http://scholarspace.manoa.hawaii.edu/bitstream/handle/10125/4577/crippenrobinson.pdf?sequence=1>>

DAVIES, MARK and MICHAEL FERREIRA. (2006-) **Corpus do Português**: 45 million words, 1300s-1900s. Disponível online em <<http://www.corpusdoportugues.org>>

DOS SANTOS, M. S. **Direito autoral na era digital**. Dissertação de mestrado em direito. Pontifícia Universidade Católica de São Paulo. 2008.

DUDIS, P. MATHUR, G. e MIRUS, G. 2009. **Bringing a corpus in line with deaf communities**. In Proceedings from the Sign Lan-

guage Corpora: Linguistic Issues workshop, University College of London.

DWYER, A.M. (2006). **Ethics and Practicalities of Cooperative Fieldwork and Analysis**. In J. Gippert, N.P. Himmelmann, & U. Mosel (Eds.), *Essentials of Language Documentation*. Berlin: Mouton de Gruyter, pp. 31--66. Disponível para download em <http://anthroweb.ucsd.edu/~jhaviland/AudVid/AudVidReadings/GippertEtAl/LangDoc_02_Dwyer.pdf>

FELIPE, T. A. **Libras em Contexto** : Curso Básico : Livro do Estudante / Tanya A. Felipe. 8ª. edição- Rio de Janeiro : WalPrint Gráfica e Editora, 2007. Disponível para download em <<http://www.librasemcontexto.org/>>

FESTA, P. S.V. **Youtube e surdez: análise de discursos de surdos no Ambiente virtual**. 2012. 170 P. Dissertação (Mestrado em Distúrbios da Comunicação) Universidade Tuiuti do Paraná, Curitiba, Paraná, 2012.

FURUNO, F. **A Evolução das mídias e a internet: Cultura Participativa transformando os meios de comunicação**. Dissertação de Mestrado. Mestrado em Comunicação. Universidade Anhembí. São Paulo. 2010

GAURAV, M.; RATHMANN, C. **Variability in verbal agreement forms across four signed languages**. In: Goldstein, L.; WHALEN, D.; BEST, C. (Orgs.). *Laboratory Phonology 8*. Berlin: Mouton de Gruyter, 2006, p. 289-316

GERALDI, J.W. **Deslocamentos no ensino de objetos a práticas; de práticas a objetos**. *Revista Línguas & Letras*. Universidade Estadual do Oeste do Paraná. 2010

HIDDINGA, A. e CRASBORN, O. **Signed languages and globalization**. *Language in Society*, 40, pp 483-505.2011 doi:10.1017/S0047404511000480.

HOCHGESANG, J.A., VILLANUEVA, P.P. MATHUR,G. LILLO-MARTIN, D. **Building a Database while Considering Research Ethics in Sign Language Communities, 4th Workshop on the Representation and Processing of Sign Languages: Corpora and**

Sign Language Technologies. 112-115
<http://homepages.uconn.edu/~dcl02005/dlm/publications_files/hoc_hgesang_et_al_2010_lrec.pdf>

HOWARD, P. N. **The Arab Spring's Cascading Effects**, Fevereiro de 2011, disponível em <<http://www.psmag.com/navigation/politics-and-law/the-cascading-effects-of-the-arab-spring-28575/>>

IBBOTSON P (2013) **O escopo da teoria baseada no uso**. frente. Psychol. 4 : 255. doi:10.3389/fpsyg.2013.00255. disponível para download em <<http://www.frontiersin.org/journal/downloadfile.ashx?pdf=1&fileid=42371&articleid=46946&version=1&contenttypeid=21&filename=fpsyg-04-00255.pdf>>

JOHNSTON, T. A. 2004. **W(h)ither the Deaf Community? Population, Genetics, and the Future of Australian Sign Language**. American Annals of the Deaf 148: 358–75 republicado em Sign Language Studies 6.2 (2006): 137-173

KILGARRIFF, A. & GREFENSTETTE, G. (2003) **'Introduction' to 'The Web as Corpus'**, special issue of Computational Linguistics 29(3): 333–47 Reprinted in Practical Lexicography: a Reader. Fontenelle, editor. Oxford University Press. 2008

_____**Web as Corpus**. Em: RAYSON, P. et al (orgs.). UCREL Technical Papers: 13. Proceedings of Corpus Linguistics. 2001 Conference. University of Lancaster, Lancaster, UK, 2001. p.342-44. Disponível para download em <<http://ucrel.lancs.ac.uk/publications/cl2003/cl2001%20conference/papers/kilgarri.pdf>>

KOESTER, A. **Building small specialised corpora** Em The Routledge Handbook of Corpus Linguistics. O'Keeffe A. and McCarthy, M. (Ed.) Routledge. New York, NY. 2010

KONRAD (2011) **Sign Language Corpora Survey**. Em The Lexical Structure of German Sign Language (DGS) in the Light of Empirical LSP Lexicography on how to integrate iconicity in a corpus-based lexicon model. Doctoral Thesis. University of Hamburg.

Tubingen: Narr Verlag. 2011 disponível em <http://www.sign-lang.uni-hamburg.de/dgs-korpus/tl_files/inhalt_pdf/sl-corpora-survey_update_2012.pdf>

LANE, H.; HOFFMEISTER, R.; BAHAN, B. **A journey into the deaf-word**. San Diego, CA; Dawn Sign Press, 1996

LEE, D. Y. W. **What corpora are available?**. em M. McCarthy & A. O'Keefe (eds), *The Routledge Handbook of Corpus Linguistics*, Routledge, Abingdon. pp. 107-121. 2010

LEITE, Tarcisio de Arantes. **A segmentação da língua de sinais brasileira (libras): Um estudo lingüístico descritivo a partir da conversação espontânea entre surdos**. 2008. Tese de Doutorado. Universidade de São Paulo.

LYONS, John. **Linguagem e Linguística** : uma introdução. Rio de Janeiro: LTC, 1987. 322p

MACHADO, F.A. **Simetria na poética visual na língua de sinais brasileira**, Dissertação de mestrado, Universidade Federal de Santa Catarina. Florianópolis, 2013

MAGNANI, J.G.C. **“Vai ter música?”**: para uma antropologia das festas juninas de surdos na cidade de São Paulo. Pontourbe, revista do núcleo de antropologia urbana da USP. Disponível em <<http://www.n-a-u.org/magnani1-2007.html>>

MARCUSCHI, L. A. **Gêneros Textuais**_Definições e Funcionalidade. Em *Gêneros Textuais e ensino*. Dionísio, A.P., Machado, A.R. e Bezerra, M.A.(org.) 5 ed. Rio de Janeiro. Lucerna. 2007

MARTELOTTA, M. E. (org.). **Manual de Linguística**. São Paulo: Contexto, 2010

McENERY, T. and WILSON, A. **Corpus linguistics, An Introduction**. 2 ed, 2001. Edinburgh University Press

MELLO, S. F. M. **Comunicação e organizações na sociedade em rede: novas tensões, mediações e paradigmas**. 2010. 271 fl. Dissertação (Mestrado em Ciências da Comunicação) - Universidade de

São Paulo, Escola de Comunicações e Artes, São Paulo. Disponível para download em <<http://www.teses.usp.br/teses/disponiveis/27/27154/tde-17082011-110313/publico/selmaferrazmottamello.pdf>>

MESSINEO, C. 2008. **Fieldwork and documentation of speech genres in indigenous communities of Gran Chaco**: Theoretical and methodological issues. *Language Documentation & Conservation* 2(2):275–295

NEWMeyer, F.J. **Grammar is grammar and usage is usage**. University of Washington, Seattle, WA. 2003

PADDEN, CAROL "The deaf community and the culture of deaf people". In: WILCOX, S. (Ed.) *American Deaf Culture: an anthology*. Burtonsville, MD: Lindtok Press, 1989

_____. **Deaf in America: Voices from a culture**. Carol Padden & Tom Humphries. 1988. Harvard University Press. Cambridge Massachusetts and London, England. Disponível para visualização parcial em <http://books.google.com.br/books?id=_icCgQ5S06wC&lpg=P1&dq=padden&hl=pt-BR&pg=PP1#v=onepage&q=padden&f=false>

_____. Padden, C. & Humphries, T. **Inside deaf culture**, Harvard University Press, Cambridge, Massachusetts London, England 2005

POSSENTI, S. **Os humores da língua: análises lingüísticas de piadas**. 3ª ed. Campinas: Mercado de Letras, 2002

REPPEN, R. **Building a corpus**: what are the key considerations? Em *The Routledge Handbook of Corpus Linguistics*. O'Keeffe A. and McCarthy, M. (Ed.) Routledge. New York, NY. 2010

Revista da Feneis, nº 44, pág. 22.

SERRANO, P.H.S.M. **Cognição e interacionalidade através do YouTube**. Biblioteca On-line de Ciências da Comunicação, 2009

disponível para download em <<http://www.paulohsms.com/wp-content/uploads/2011/03/serrano-paulo-cognicao-interacionalidade-YouTube.pdf>>

SERRANO, PHSM; PAIVA, Cláudio Cardoso. **A Nova (Des) Ordem da Cibercultura Normas de Uso, Restrição e Censura no YouTube**. In: CONGRESSO BRASILEIRO DE CIÊNCIAS DA COMUNICAÇÃO. 2008

SILVEIRA, G.L. e AMARAL, M.F. **Movimento surdo e o ciberativismo através do YouTube e do Facebook**. Trabalho apresentado no IJ05 – Comunicação Multimídia do XIII Congresso de Ciências da Comunicação na Região Sul realizado de 31 de maio a 2 de junho de 2012

SCHALLENBERGER, Augusto. **Ciberhumor nas comunidades surdas**. 2010, 75 f. Dissertação (Mestrado) – Universidade Federal do Rio Grande do Sul. Faculdade de Educação. Programa de Pós-Graduação em Educação. Porto Alegre: 2010.

SCHEMBRI, A. e JOHNSTON, T. **Usage based grammars and sign languages**. La Trobe University, 2011

SCHNEUWLY, B. **Gêneros e tipos de discurso: considerações psicológicas e ontogenéticas**, In. / tradução e organização ROJO, R.; CORDEIRO, G. S., *Gêneros orais e escritos na escola*, Campinas, SP: Mercado das Letras, 2004.

SLOETJES, Han; WITTENBURG, Peter; SOMASUNDARAM, Aarthi. **ELAN-Aspects of Interoperability and Functionality**. In: INTERSPEECH. 2011. p. 3249-3252.

THOMPSON, P. **Building a specialised audio-visual corpus** Em *The Routledge Handbook of Corpus Linguistics*. O’Keeffe A. and McCarthy, M. (Ed.) Routledge. New York, NY. 2010

TORRENT, Tiago Timponi Torrent. **Usage-based models in linguistics: an interview with Joan Bybee**. Revista Linguística / Revista do Programa de Pós-Graduação em Linguística da Universidade Federal do Rio de Janeiro. Volume 8, número 1, junho de 2012.

ISSN 1808-835X 1. <[http://www.lettras.ufrj.br/poslinguistica/ revis-talinguistica](http://www.lettras.ufrj.br/poslinguistica/revis-talinguistica)>

WILCOX, Sherman; WILCOX, Phyllis Perrin. Aprender a ver. Petrópolis, RJ, Editora Arara Azul, 2005

APÊNDICE

1. Declaração de ética**Declaração de ética da sociedade de linguística da língua de sinais (SLLS) para a pesquisa com língua de sinais****Julho de 2013**

Já foram estabelecidas diretrizes éticas gerais para os pesquisadores em linguística e abrangem a maior parte das questões gerais necessárias para a pesquisa de língua de sinais (ver declaração de ética da LSA). Esta declaração não pretende duplicar tais documentos, mas sim visa a estabelecer um panorama geral que complementa essas diretrizes éticas gerais ao incluir considerações éticas mais específicas relativas a sinalizantes e comunidades surdas. Usuários de língua de sinais e as comunidades têm sido tradicionalmente marginalizados e os pesquisadores devem estar sempre cientes de que isso pode resultar em desigualdades nas relações de poder no que diz respeito a informantes da língua de sinais.

Pesquisa com línguas de sinais, porém, são feitas sob muitas formas e acontecem em muitas situações diferentes. Tomemos, por exemplo, as radicalmente diferentes configurações envolvidas no trabalho com língua de sinais na zona rural/interior/pequenas comunidades comparado com o trabalho com comunidades surdas urbanas/nacionais. Além disso, a pesquisa linguística de línguas de sinais é realizada a partir de uma variedade de subdisciplinas com métodos e objetivos específicos. Este documento fornece orientações gerais que precisam ser adaptada a cada situação particular de pesquisa. Exigências éticas mais rigorosas que não são tratadas aqui precisam ser criadas, por exemplo, para a pesquisa com as crianças como as pesquisas que abordam a aquisição de língua de sinais ou no trabalho em deficiências de língua de sinais em sinalizantes. Mais diretrizes para abordar tais situações e metodologias de investigação específicos serão encomendadas ao Comitê de Ética da SLLS.

As considerações éticas avaliadas nesta Declaração devem ser vistas como guia permanente durante todo o processo de investigação

1 – Responsabilidades para com os informantes individualmente

Os participantes da pesquisa que atuam como informantes compartilham seus conhecimentos e outros aspectos de suas vidas com o pesquisador, e podem referir-se à informações de identificação pessoal de si mesmo ou de terceiros durante a coleta de dados. Estes aspectos

requerem uma consideração cuidadosa das questões de privacidade. Os linguistas que trabalham com línguas de sinais, quer sejam Surdos ou ouvintes, também devem levar em conta a dimensão adicional inevitável da gravação em vídeo que torna impossível separar os dados da identidade pessoal do participante da pesquisa. Os direitos e desejos dos sinalizantes com relação a tais gravações devem ser respeitados em todos os momentos. Para este fim, o pesquisador deve cumprir os seguintes requisitos básicos:

- O linguista deve explicar de uma forma simples e acessível a que os objetivos gerais a pesquisa é destinada antes do início da coleta de dados (e, quando for o caso, os demais objetivos específicos da pesquisa após a coleta de dados estar concluída*). Se o linguista não é proficiente ou suficiente na língua de sinais do informante, ele/ela deve pedir a mediação linguística por um intérprete profissional ou por outro pesquisador ou informante que é proficiente na língua de sinais em questão.

- Antes de começar a trabalhar, o pesquisador deve pedir o consentimento formal do informante, quer na forma escrita na língua falada do local quer na sua língua de sinais. A primeira opção depende do nível de alfabetização do informante na língua falada local, mas verifica-se que isso cria uma relação assimétrica desde o começo da pesquisa, por isso, devem ser oferecidas as duas possibilidades de leitura do termo ao informante. O consentimento informado não deve ser visto como uma formalidade, mas como uma maneira de tornar o consultor ciente de todas as implicações do fornecimento de dados, da cessão da imagem para filmagem, e (quando aplicável) do arquivamento e compartilhamento desses dados a longo prazo. Para este fim, o consentimento informado deve ser explícito sobre a configuração, duração e número de sessões de pesquisa, o eventual uso e armazenamento dos dados, quando gravados, a possibilidade dos dados serem ainda analisados por outros pesquisadores da equipe, serem incluídos em publicações ou em apresentações em eventos de pesquisa e de serem incorporados em repositórios de dados que podem ter acesso restrito ou aberto. Deve ser dada atenção especial a um tópico importante: a disseminação dos dados em acesso aberto na internet. Isso porque a proteção da identidade do sinalizante em um vídeo específico é impossível garantir nesse caso. Para preservar a confidencialidade do sinalizante, ele/ela também deve se perguntar se ele / ela aceita ter seu / suas informações de metadados associadas aos arquivos de dados, bem como ser por sua contribuição para a investigação através do registro do seu nome nesses dados. Uma cópia do termo de consentimento informado (a forma escrita ou a gravação em língua de sinais) devem ser disponibilizados para o informante.

O termo de consentimento informado é uma condição para que a pesquisa seja aprovada por um comitê de ética e essa aprovação tem sido exigida por periódicos também.

- Os investigadores devem sempre avaliar se alguma espécie de remuneração ou compensação é adequada para o tempo e o esforço do informante.

- Os resultados da investigação e, sempre que possível, cópias dos dados em si devem sempre ser feita e disponibilizadas para os informantes que participaram da pesquisa.

- Os pesquisadores, Surdos ou ouvintes, devem esforçar-se para treinar os surdos participantes da pesquisa como assistentes ou colegas de pesquisa, na medida em que é estabelecida uma relação de troca de longo prazo

2. Responsabilidade para com as comunidades de surdos

Trabalhar com participantes de pesquisa individuais não tem relação com a comunidade surda que eles pertencem. Os linguistas não devem sempre tomar como certo que o seu trabalho é, por definição, benéfico para as comunidades de sinalizantes e devem tentar explicar de forma clara qual a vantagem que a pesquisa linguística que tem como alvo a língua de sinais pode trazer para eles. Visibilidade e reconhecimento das línguas de sinais como línguas naturais humanas deve ser uma das metas básicas para a pesquisa linguística realizada com a comunidade surda.

O pesquisador deve estar familiarizado com as normas culturais e valores da comunidade surda em questão e respeitá-los, especialmente se ele/ela não é um membro dessa comunidade. O pesquisador deve sempre se esforçar para se comunicar diretamente com os informantes por meio da língua de sinais alvo e mostrar atitudes positivas para a comunidade surda. Ele/ela deve estabelecer uma forte relação com a comunidade antes de começar seu trabalho, e, se for ouvinte, ele/ela devem trabalhar em estreita colaboração com os assistentes ou colegas surdos, de preferência acadêmicos.

Na medida em que a comunidade em questão já tenha organizado associações representativas, o linguista deve discutir com eles o acesso, arquivamento e distribuição dos dados e resultados da investigação antes do início do projeto. Além disso, é aconselhável discutir com a comunidade surda quais aplicações da pesquisa consideram uma prioridade ou urgência, de modo a tentar atender a essas necessidades na medida do possível.

Os resultados da pesquisa devem ser disponibilizados para a comunidade surda através da língua de sinais que é alvo da pesquisa, bem como através da forma escrita da língua falada do local. Publicação em Inglês e Sinais Internacionais devem ser considerada como algo desejável, de modo a promover o trabalho entre diversas comunidades e línguas.

De acordo com o contexto nacional, os critérios explícitos para uma investigação eticamente adequada na língua de sinais devem ser considerados na avaliação das propostas de pesquisa por organizações que financiam a pesquisa acadêmica.

Respeito à variação lingüística individual dentro da comunidade surda é importante, dadas as diversificadas biografias linguísticas dos indivíduos que a formam.

3. Responsabilidade para com os estudantes e com o público em geral

Linguistas das línguas de sinais devem fazer um esforço para tornar os resultados de pesquisa sobre as línguas de sinais visíveis no meio acadêmico em geral e nos subcampos específicos onde eles são diretamente relevantes.

Linguistas das línguas de sinais igualmente devem se esforçar para tornar seus resultados de pesquisa disponíveis para o público mais amplo, não-especializado, tendo como objetivo final generalizar o conhecimento sobre língua de sinais e sinalizantes para acabar com os preconceitos tradicionais sobre eles.

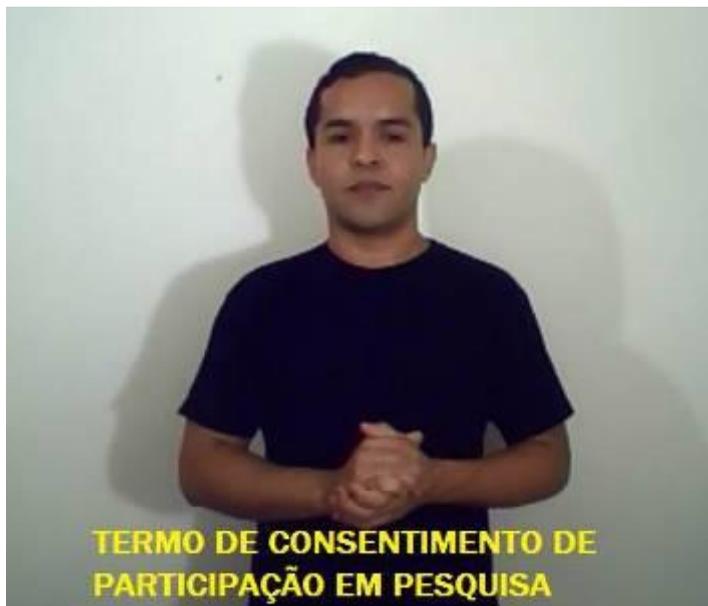
4. Coleguismo e boas práticas

Dado o histórico de desequilíbrio entre surdos e pesquisadores ouvintes no meio acadêmico e ambientes de pesquisa em geral, o uso de uma língua de sinais deve ser considerado como a forma padrão de comunicação pois promove a igualdade, quer no local de trabalho ou em reuniões científicas, bem como em trocas informais. O uso da língua de sinais é a única abordagem verdadeiramente inclusiva para compartilhamento de informações e, como tal, deveria constituir as melhores práticas entre colegas surdos e ouvintes.

***n.t. subtende-se quando a explicação detalhada anteriormente puder enviesar o processo de coleta de dados.**

ANEXOS

Anexo 1 – Termo de consentimento de participação em pesquisa



1. Termo de consentimento de participação em pesquisa

Do que se trata essa pesquisa?

Essa é uma pesquisa de mestrado que tem como finalidade a obtenção do título de mestre em linguística pelo autor da pesquisa. Como tema de pesquisa o mestrando escolheu discutir a questão de como é feita a coleta de dados em vídeos na internet para a formação de um corpus linguístico de língua de sinais. Para isso pré-selecionou alguns

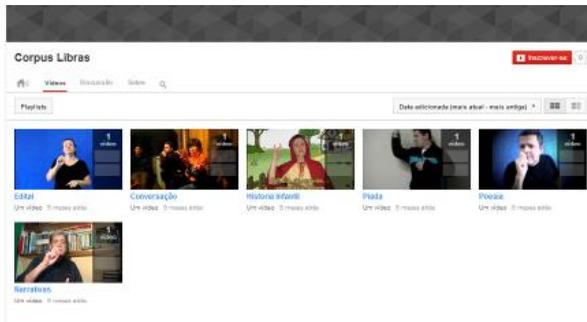
vídeos na internet, incluindo o seu disponibilizado na plataforma youtube.com para postar em um canal próprio na mesma plataforma.

O que é um corpus linguístico?

Corpus linguístico é qualquer conjunto de textos escritos, falados ou sinalizados que são organizados por alguém para ser analisado. Em geral, os corpora linguísticos são analisados por linguistas que são pesquisadores que trabalham em universidades, locais de produção de conhecimento científico. Esses profissionais dedicam seus trabalhos a pesquisar e estudar aspectos de uma ou mais línguas de forma a produzir conhecimento sobre elas.

Como serão utilizados meus dados?

Os vídeos selecionados que você postou em seu canal, ou que alguém postou com o seu conteúdo será repostado no canal do pesquisador na plataforma youtube que tem acesso livre para qualquer pessoa com acesso a internet. Nessa plataforma os vídeos serão organizados por gênero, que é uma classificação para uma característica específica de uma forma de usar a língua, por exemplo: discurso, notícia, entrevista, conversa etc...



Que tipo de informações aparecerão nas descrições do vídeo?

Na descrição dos vídeos aparecerão as seguintes informações:

Nome da pessoa(s) que aparece no vídeo.

Sexo:

Idade que tinha na filmagem desse vídeo:

Local onde mora:

Local onde morou nos últimos 10 anos:

Comunidades de surdos com quem conviveu nos últimos 10 anos (para o caso de ter morado por pouco tempo em algum outro lugar ou ir com certa frequência a uma comunidade de surdos distante da sua região)

Com que idade aprendeu língua de sinais:

Que comunidade de surdos que conviveu para aprender a língua de sinais?

Familiares surdos?

Participa de comunidades restritas de surdos? (Religiosas, acadêmicas, associação de surdos, teatro etc)

Conhece ou utiliza outras línguas de sinais?

Você pode optar por retirar qualquer uma dessas informações, por exemplo, omitir a idade ou o fato de conhecer outras línguas de sinais.

O que eu ganho com isso?

A pesquisa científica com língua de sinais tem como primeiro objetivo promover o reconhecimento da língua de sinais na sociedade e no meio acadêmico, resultando em melhores condições de vida para toda a comunidade surda, através de políticas linguísticas que favoreçam o reconhecimento e uso da língua de sinais pelos indivíduos surdos. O corpus da pesquisa tem como objetivo secundário possibilitar a pesquisa da língua por outros pesquisadores de forma que sua produção linguística poderá figurar em vários trabalhos de pesquisa. Como participante individual você receberá uma versão da pesquisa em língua de sinais através de um DVD que lhe dará acesso ao conhecimento produzido nessa pesquisa. Caso queira os vídeos podem informar o link da sua postagem original aumentando o acesso ao seu canal na plataforma youtube o que pode gerar renda para seu canal através dos programas de monetização da própria plataforma além de divulgar seus vídeos o que é uma ótima ferramenta se o seu canal também divulga trabalhos individuais seus como palestras, oficinas, apresentações teatrais etc.

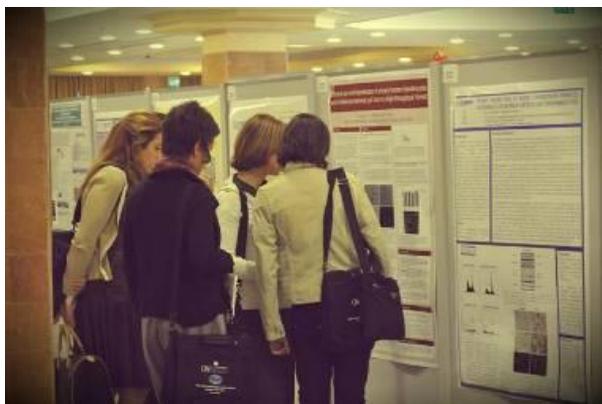
O que o pesquisador ganha com isso?

Através da pesquisa científica os pesquisadores obtêm títulos outorgados pelas instituições de ensino que lhes habilitam a exercer funções acadêmicas. Os títulos não são resultados diretos do corpus em si, mas da pesquisa científica desenvolvida a partir deles, que são elementos de livre escolha do pesquisador. Em hipótese nenhuma haverá uso comercial dos vídeos. Os materiais de pesquisa que se utilizam dos vídeos não serão vendidos, comercializados ou utilizados como permuta comercial. Os usos decorrentes do material como artigos científicos, publicações, pesquisas etc também não poderão ter qualquer fim comercial. Haverá uma orientação explícita sobre isso inserida antes da exibição do vídeo no canal.

Quem poderá acessar meus vídeos? Como eles serão utilizados?

Os vídeos serão acessados pelas mesmas pessoas que acessam o vídeo no seu canal atualmente, o público em geral que tem acesso a internet. A diferença é que no canal do pesquisador acredita-se que haverá um fluxo maior de visitantes também pesquisadores já que a divulgação do canal será feita em trabalhos de pesquisa científica. Esses pesquisadores por sua vez, possivelmente se interessarão no uso dos vídeos para realização de novas pesquisas. Dessa forma o vídeo, imagens do vídeo e referência aos dados contidos neles poderão figurar em outros meios de pesquisa como:

Pôsteres de trabalhos acadêmicos apresentados em congressos e eventos científicos



Palestras em eventos científicos e acadêmicos



Uso para ensino em sala de aula



Quem tem o controle do acesso do vídeo?

O canal em que o vídeo será respondado é de controle do pesquisador. Só ele pode bloquear, retirar ou monitorar o acesso do vídeo. Os vídeos selecionados serão postados por tempo indeterminado. Caso em algum momento você deseje retirar o vídeo do ar deverá enviar uma solicitação por email ou correio para o pesquisador cujos dados estão abaixo:

Edgar Correa Veras – CPF 51466295287 – RG 4077690 SSP-PA

E-mail 1 – **EDGAR.CORREA@UFSC.BR**

E-mail 2 – **EDGAR.UFSC@GMAIL.COM**

Endereço comercial – Universidade Federal de Santa Catarina
– Centro de Comunicação e Expressão - Coordenadoria de intérpretes
de libras – Campus Universitário Reitor João David Ferreira Lima, s/n,
Trindade, Florianópolis - Santa Catarina, 88040-900

Endereço residencial – Rua João Pio Duarte Silva, nº 844, Apt
302B, Córrego Grande, Florianópolis – SC – CEP 88037-001

Quais vídeos meus serão postados?

Os vídeos postados são os descritos abaixo.

Endereço eletrônico do vídeo	Título do vídeo no youtube	Frame do vídeo	Resumo do conteúdo	Tempo de duração do vídeo

Serão inseridas informações de utilização do vídeo antes do início do mesmo e poderão ser retiradas informações do vídeo que exponham terceiros ou que sejam consideradas sigilosas ou constrangedoras. O vídeo editado será postado para visualização restrita do seu proprietário antes de ser disponibilizado para acesso ao público em geral

ANEXO 2 – Questionário da Pesquisa



2. Questionário de pesquisa – Corpus da língua de sinais brasileira e a comunidade surda

Primeira parte – Questionário de sondagem individual

Nome:

Idade:

Funções que já exerceu nos movimentos surdos? Fale um pouco sobre esse trabalho.

Com quantos anos aprendeu língua de sinais?

Foi alfabetizado em língua portuguesa?

Que língua(s) utiliza para comunicação com ouvintes que não sabem língua de sinais.

Língua de sinais apenas (só se comunica com ouvintes sinalizadores)

Intérprete de libras Gestos Português escrito Português oral

Como você considera sua proficiência em libras?

Ruim, tenho dificuldade de expressar todo meu pensamento em libras

Boa, consigo expressar o que necessito em libras

Ótima, sou muito fluente em libras

Excelente, tenho ótima performance em todos

<p>os gêneros em libras (piada, poesia, narrativas etc)</p> <p>Ao que você atribui sua proficiência em libras?</p> <p>Convivência com outros surdos</p> <p>Ter sido estimulado desde criança em libras</p> <p>Ter pais ou parentes surdos</p> <p>Assistir vídeos e outros materiais com gêneros específicos em libras</p> <p>Qual a primeira comunidade surda com a qual você teve contato? Ela era formada por surdos nativos de qual estado/região?</p> <p>Quantos anos você conviveu nessa comunidade?</p> <p>Com qual comunidade surda você mais conviveu nos últimos 10 anos?</p>
<p>2 – Apresentação de corpus e línguas de sinais</p>
<p>Corpus linguístico compreende toda coletânea de dados que representem o uso de uma língua. Na pesquisa científica em geral se determinam critérios para formação desse corpus que tende a atender uma determinada finalidade de pesquisa, como a investigação de um elemento específico do sistema linguístico. No entanto um corpus linguístico pode servir a diversos meios como aqueles elaborados por dicionários, que buscam apresentar os ítems lexicais e seus significados ou mesmo o conjunto de obras de uma biblioteca, que buscam preservar poesias, crônicas e diversas outras obras literárias que representam gêneros da língua escrita. As línguas de sinais demandam registros em vídeo para formação de seu corpus, ou menos comum, o registro em escrita de sinais.</p>
<p>3 – Questões sobre o uso de corpus da língua</p>
<p>Você já recorreu a corpus da língua de sinais? Se sim, qual a finalidade?</p> <p>Você reconhece nesse corpus a representatividade da língua brasileira de sinais para os quais ele se propõe? Você teria algum comentário para fazer sobre esses trabalhos?</p> <p>Você acha que é importante manter registros da língua de sinais? Qual a importância desse corpus na</p>

sua opinião?
4 – Questões sobre a utilidade da pesquisa linguística para a comunidade surda
<p>Você conhece as pesquisas que tem sido desenvolvida na academia científica brasileira com a língua de sinais?</p> <p>Qual a importância desse trabalho?</p> <p>Você acha que outros tipos de pesquisa poderiam ser desenvolvidas?</p> <p>Você acredita que as pesquisas desenvolvidas atualmente vão de encontro as necessidades da comunidade surda?</p>
5 – Questões sobre corpus e gênero linguístico
<p>5 – Você poderia citar alguns tipos específicos de textos que você conhece em língua de sinais?</p> <p>Em que contexto eles são utilizados? Com qual finalidade? Que elementos linguísticos da libras você acredita marcar cada um desses gêneros?</p> <p>Que tipos de textos você acredita que devam compor um corpus representativo da língua de sinais?</p> <p>Você acredita que há tipos de textos em língua de sinais que são mais usados do que em língua portuguesa?</p> <p>Você acredita que existam tipos de textos usados apenas em língua de sinais?</p>

ANEXO 3

Análise dos 100 primeiros vídeos resultantes da busca por “li-bras” na plataforma youtube

	Vídeo	Categoria	Descrição do vídeo	Link do vídeo	Língua utilizada	Personagens	Objetivo do Vídeo
1	Curso Básico de Libras Vídeo Institucional INES	Educação	No dia-a-dia nós pessoas que não somos portadoras de nenhuma deficiência vivemos num mundinho fechado para os problemas dos outros. (...)	http://www.youtube.com/watch?v=WDAOH_5pN1g	Libras com legendas e voz em Português	Surdos	Ensinar libras para pessoas ouvintes
2	Curso de Libras - Módulo Prático: Aula 1 - Saudações e Apresentação	Educação	Vídeo feito para o projeto Educopédia e faz parte de um curso criado para auxiliar os professores que trabalham na SME do Rio de Janeiro.	http://www.youtube.com/watch?v=Igej4uP7ogI	Libras com legendas em Português	Surdos	Ensinar libras para professores da rede de educação do Rio de Janeiro
3	Aprenda LIBRAS com eficiência e rapidez (vocabulários básicos)	Educação	Vocabulários da LIBRAS básico no LIVRO Aprenda LIBRAS com eficiência e rapidez Volume 1 e 2.	http://www.youtube.com/watch?v=III-BelLxSUW9k	Libras com legenda	Surdo	Apresentar glossário da libras para auxiliar aprendizes ouvintes de libras

4	Curso Instrutor de Libras - Parâmetros e Classificadores - Aula 1 - Prof. Luiz Albérico Falcão	Educação	A leitura de imagem como exercício dialógico de simulação em Libras tem sido uma estratégia construtiva e elaborativa eficiente. Para tanto, é necessário exercício crítico e reflexivo sobre o uso dos recursos disponíveis e a criação de novas estratégias comunicacionais e educacionais.	http://www.youtube.com/watch?v=1BFDVig2HOA	Língua portuguesa oral com exemplos em libras	Ouvinte conhecedor dos sinais apresentados	Orientar quanto a aspectos linguísticos da libras para alunos ouvintes
5	Verbos na libras	Educação	Sem descrição	http://www.youtube.com/watch?v=...	Libras com legenda em	?	Apresentar glossário de verbos na libras
6	Alexandre Elias - VERBOS EM LIBRAS Legendado	Educação	Neste vídeo o Ms. Educador de surdos Alexandre Elias apresenta os 110 Verbos mais utilizados pela comunidade surda.(...) Espero que esse material possa auxiliar a todos os admiradores desta língua, educandos ouvintes envolvidos e surdos	http://www.youtube.com/watch?v=BuAZCJHOW10	Libras com legenda em português	Ouvinte conhecedor dos sinais apresentados	Auxiliar aprendizes ouvintes de libras e surdos

7	Lição 1 LIBRAS Língua Brasileira de Sinais	Pessoas e blogs	Curso básico	http://www.youtube.com/watch?v=OdakxBs2Ik	Libras com legendas do	Ouvintes conhecedores	Auxiliar aprendizes ouvintes
8	BE - LI - BRAS - mostrando os sinais que ele já sabe!	Entretenimento	Be, 1 ano e meio, na hora do café com o papai, Lincoln, que é surdo! Estamos na casa da vovó, e queremos ver se ele mostra os sinais que ele já sabe, pois aprendeu conversando com o pai, com os amigos e parentes surdos!! Ele não é lindo? (...)	http://www.youtube.com/watch?v=OdakxBs2Ik	Libras e língua portuguesa oral	Criança Coda interagindo com mãe ouvinte e pai surdo	Apresentar a proficiência de uma criança Coda em língua de sinais
9	manifestação neste sábado! com áudio	Notícias e política	Manifestação do Movimento Surdo sim,calado não! No próximo sábado dia 6 de julho de 2013. Para mais informações curta a fan-page https://www.facebook.com/SurdoSimCala..	http://www.youtube.com/watch?v=z5NGj_PbPKc&feature=c4-over	Libras com interpretação em língua portuguesa oral	Surdo	Divulgar atividades políticas e orientar a comunidade surda politicamente

10	Filme em Libras "Resgate no silêncio" (com legenda)	Filmes e desenhos	Filme "resgate no silêncio" em Libras, mostra a falta de acessibilidade dos surdos em órgãos públicos OBS: Os logos que aparecem no início do vídeo das produtoras MGM e a Columbia (sony) são meramente obras de ficção não tendo o vídeo nenhum vínculo com as mesmas.	http://www.youtube.com/watch?v=ZXKdP7uRs0	Libras (Embora o título informe que há legendas, o que talvez mostrasse o desejo de fazer a denúncia a surdos e	Ouvintes aprendizes de libras	Conscientizar/Denunciar a falta de acessibilidade dos surdos em órgãos públicos.
11	em libras, mae e pais e filho	Automóveis	em libras que é família	http://www.youtube.com/watch?v=...	Libras	Surdos	Entreter surdos com vídeo de comédia em libras
12	Fico assim sem você - LIBRAS	Educação	Exercício da turma de teatro III da Fazenda Arte usando a Língua Brasileira de Sinais - LIBRAS. Direção de Valeska Picado	http://www.youtube.com/watch?v=nZW7C1neXOM	Libras	Aprendizes ouvintes	Apresentar conteúdo aprendido em curso de libras
13	Gramática De Libras	Educação	<i>Nenhuma descrição disponível.</i>	http://www.youtube.com/watch?v=...	Libras	Surdo	Ensinar aprendizes ouvintes quanto à gramática da libras

14	Dicas de Convivência - Libras + Áudio descrição	Sem fins lucrativos/ativismo	O Instituto Mara Gabrielli (IMG) é uma organização sem fins lucrativos que desenvolve e executa projetos que contribuem para a melhoria da qualidade de vida de pessoas com deficiência. (...)	http://www.youtube.com/watch?v=KWzHZZUc20	Língua portuguesa oral com janelas de libras	Ouvintes	Conscientizar o público em geral quanto ao comportamento frente à pessoas deficientes
15	Curso básico de Libras II 000	Educação	<i>Nenhuma descrição disponível.</i>	http://www.youtube.com/watch?v=...	Libras com legenda e	Surdos	Ensinar Libras para aprendizes ouvintes
16	Alfabeto em LIBRAS	Educação	Letras do Alfabeto mostradas em Língua Brasileira de Sinais	http://www.youtube.com/watch?v=...	Libras com legenda em	-	Ensinar Libras para aprendizes ouvintes
17	LIBRAS - SENTIMENTOS	Pessoas e blogs	Sinais de vários sentimentos e descrição de personalidade.	http://www.youtube.com/watch?v=MaOpziW7P-U	Língua portuguesa oral com exemplos em libras	Deficiente auditiva com alunos ouvintes	Registrar parte de uma aula de libras para divulgação do curso, aprendizado da libras e constituição de material de estudo

18	Teatro Bullying em Libras	Educação	Vídeo produzido por profissionais surdos e ouvintes do CEADA em Campo Grande-MS, com o objetivo de proporcionar aos alunos surdos e comunidade surda o acesso às informações sobre o tema tão comentando nestes dias o Bullying.	http://www.youtube.com/watch?v=FRFwRO7Zohc	Libras	Surdos e ouvintes sinalizadores	Proporcionar acesso à informações sobre tema da atualidade para a comunidade surda.
19	LIBRAS * Mãe surda e Filha ouvinte	Educação	<i>Nenhuma descrição disponível.</i>	http://www.youtube.com/watch?v=FRFwRO7Zohc	Libras	Mãe surda e filha ouvinte	Apresentar a proficiência em libras de criança ouvinte

20	Libras - Alfabeto, números, cores e semana.	Educação	<p>LÍNGUA BRASILEIRA DE SINAIS</p> <p>A voz dos surdos são as mãos e os corpos que pensam, sonham e expressam. A Língua de Sinais envolve movimentos que podem parecer sem sentido para muitos, mas que significam a possibilidade de organizar as idéias, estruturar o pensamento e manifestar o significado da vida para os surdos. (...)</p>	http://www.youtube.com/watch?v=0I3NX_oQAk	Língua portuguesa oral com exemplos em libras	Ouvinte	Ensinar libras para ouvintes
21	Faroeste Caboclo Libras	Automóveis	Faroeste Caboclo - música da banda Legião Urbana traduzida para a Língua de Sinais Brasileira por Tom Min Alves, em exercício da disciplina de Laboratório III do curso Letras-Libras da UFSC	http://www.youtube.com/watch?v=R8QDHNvQYak	Libras e língua portuguesa oral (música) simultaneamente	Ouvinte	Apresentar interpretação pessoal de uma música em libras

22	LIBRAS - FRASES	Pessoas e blogs	Diversas frases para estudar verbos e palavras.	http://www.youtube.com/watch?v=08klwvjack	Língua portuguesa oral e exemplos em libras	Deficiente auditiva	Registrar parte de uma aula de libras para divulgação do curso, aprendizado da libras e constituição de material de estudo
23	Curso Libras - Aula 6 - Pronomes Possessivos e adjetivos - Prof. Luiz Albérico Falcão	Educação	Aula gravada do Curso de Libras com o objetivo de universalização e popularização da Língua Brasileira de Sinais. (...)	http://www.youtube.com/watch?v=DREM5u-52_M	Língua portuguesa oral com alguns	Ouvinte	Ensinar língua de sinais para ouvintes
24	FRASES EM LIBRAS.flv	Viagens e Eventos	Este vídeo foi desenvolvido com frases simples para que os alunos (as) do curso Letras Libras-UFPB, possam comparar a estrutura do Português e da Libras.	http://www.youtube.com/watch?v=iHrOnKKvfiQ	Libras com legenda em português	Ouvintes	Ensinar alunos ouvintes quanto à diferença de estrutura libras - português
25	Em libras piada " Mãe e filhos " NOVOO!	Pessoas e blogs	<i>Nenhuma descrição disponível</i>	http://www.youtube.com/watch?v=iHrOnKKvfiQ	Libras	Surdo	Entreter surdos

26	BE - LIBRAS - Mais sinais, especial pra Tia J-U..	Entretimento	Eu, papai Lincoln (LN) e a vovó Mary gostamos muito de registrar momentos com o Bernardo, sempre que viajamos nas férias brincamos com filmagens e musicas. (...)	http://www.youtube.com/watch?v=I70tpCIUNeQ	Libras e Língua portuguesa oral	Criança Coda interagindo com mãe ouvinte e pai surdo	Apresentar proficiência de criança ouvinte sinalizadora
27	Pequeno Glossário em Libras	Educação	Dividido conforma a configuração das mãos, este pequeno glossário em Libras faz parte da Formação em Libras, feita com os colaboradores do Colégio Marista Rosário, em julho de 2012, o qual fiz parte. Ministrado pela professora e tradutora Mestre em Educação Ângela Russo.	http://www.youtube.com/watch?v=hiGjcZzzPw	Libras com legenda	Aprendiz ouvinte	Registrar glossário de palavras em libras aprendidas em curso

28	Ai se eu te pego Libras	Educação	Nova versão do vídeo "Ai se eu te pego - (Michel Teló)" feito pelas alunas do curso de Fonoaudiologia da Faculdade de Ciências Médicas da SANTA CASA SP para uma avaliação da matéria de Libras.	http://www.youtube.com/watch?v=cCEhV10MFw	Libras e língua portuguesa oral (música) simultaneamente	Ouvintes	Apresentar interpretação pessoal de uma música em libras
29	O encontro - LIBRAS - parte 1	Educação	O encontro - LIBRAS - parte 1	http://www.youtube.com/watch?v=...	Libras com legendas e	Surdos	Ensinar língua de sinais para pessoas ouvintes
30	Hino 260 CCB - Sou Servo Inútil - LIBRAS * Marcelo de Brito	Música	Hino do Novo Hinarío, nº 5 da Congregação Crista no Brasil. Interpretado em LIBRAS por Marcelo de Brito	http://www.youtube.com/watch?v=0w09-SHzsZg	Libras e língua portuguesa oral	Ouvinte	Apresentar interpretação pessoal de uma música em libras
31	Me Ama - Diante do Trono - em LIBRAS	Pessoas e blogs	Interpretação feita por Lívia Martins Gomes, dia 24/05/2012 E-mail: biologia.libras@hotmail.com	http://www.youtube.com/watch?v=IcF9a9dR	Libras e língua portuguesa oral	Ouvinte	Apresentar interpretação pessoal de uma música em libras

32	Salas e Disciplinas em LIBRAS by Marcos Cunha	Educação	Galera fiz outro video para meu treinamento. https://www.facebook.com/marcos.cunha... https://twitter.com/marcosagcunha http://www.linkedin.com/home?trk=hb_t... http://marcos-cunha.blogspot.com.br/	http://www.youtube.com/watch?v=5ReFYKjplp8	Língua portuguesa oral com exemplos em libras	Ouvinte	Ensinar sinais da libras para ouvintes
33	Divulgacao em libras manifesto dia 20 de junho de 2013 Florianopolis	Pessoas e blogs	Divulgação em Libras (sem legenda) sobre os manifestos que estão ocorrendo no Brasil, e a data da manifestação em Florianópolis (dia 20 de junho as 18h, no ticen)	http://www.youtube.com/watch?v=rufC3C71IWA	Libras	Ouvinte	Divulgação para surdos de evento político
34	DM Surdo,Deaf (Libras) Explicar PortaMil Sorte Ganho voce!	Entretenimento	Bom dia! pessoas surdos união brasil estado!.você queria fazer seu cadastrar Cadastrar (...)	http://www.youtube.com/watch?v=UpikAbVXY	Libras	Surdo	Divulgação de proposta comercial para surdos

35	VIDEO AULA DE LIBRAS CPDR COM SOM	Educação	VÍDEO AULA DO PRIMEIRO MÓDULO DO CURSO DE LIBRAS, MINISTRADO EM RIBEIRÃO - PE. NO CPDR - CENTRO DE APOIO A PESSOA COM DEFICIÊNCIA DE RIBEIRÃO. PROFESSOR: DAVID SANTOS - A MAIS DE 6 ANOS NA ÁREA DE LIBRAS.	http://www.youtube.com/watch?v=E9986eFeCkI	Libras com transcrição em LP	Ouvintes aprendizes de libras	Apresentação de conteúdo aprendido em curso
36	Curso de Libras - Bolo de Laranja - Construção do cenário - Prof. Luiz Albérico Falcão	Educação	De ouvinte para ouvinte orientamos como construir o cenário de uma cozinha utilizando os recursos Marcadores e Classificadores da Libras`(...)	http://www.youtube.com/watch?v=xles-X-T_MU	Língua portuguesa oral com alguns exemplos em libras	Ouvinte	Ensinar língua de sinais para ouvintes

37	Curso de Libras - Aula 10 - Meios de comunicação e transportes - Prof. Luiz Albérico Falcão	Educação	Aula gravada do Curso de Libras com o objetivo de universalização e popularização da Língua Brasileira de Sinais. De uma forma fácil e rápida aprende-se todo o alfabeto e os números em datilologia que são algumas das configurações utilizadas no diálogo em Libras. Podem surgir variações regionais de sinais o que se deve considerar como comum e natural diante da dinâmica da língua. Grupo: libras.pernambuco@gmail.com	http://www.youtube.com/watch?v=OFI5EZxpcIg	Língua portuguesa oral com alguns exemplos em libras	Ouvinte	Ensinar língua de sinais para ouvintes
----	---	----------	---	---	--	---------	--

38	Getsêmani em Libras por Nilton Câmara (2013)	Educação	<p>Ministração da canção Getsêmani, de Leonardo Gonçalves, interpretada em LIBRAS por Nilton Câmara no Festival de Coreografia IBMC (Igreja Batista Monte Castelo) na noite de 20 de abril de 2013. Assistam e sejam abençoados !</p> <p>www.niltoncamara.net Instagram: Nilton Câmara FanPage: bo-ok.com/niltoncamaralibras</p>	http://www.youtube.com/watch?v=yOMQsQ_A1Y	Libras e língua portuguesaoral (música) simultaneamente	Ouvintes	Apresentar interpretação pessoal de uma música em libras
39	Software faz a tradução da Libras para o português - Jornal Visual	Notícias e política	<p>11/01/2013 - Para fechar a série comemorativa dos 10 anos da Língua Brasileira de Sinais, uma novidade: um software, que poderá ser instalado no telefone celular, para fazer a tradução da Libras para o português e vice versa.</p>	http://www.youtube.com/watch?v=e65LZxZTqho	Libras e língua portuguesa oral	Ouvintes intérpretes de libras e ouvintes falantes de LP	Apresentar conteúdo jornalístico para o público surdo

40	Sou Humano - Bruna Karla (LIBRAS)	Música	Não somos perfeitos e nosso Deus sabe disso! Espero que gostem das escolhas que fiz...	http://www.youtube.com/watch?v=Y2cl5NHVK	Libras e língua portuguesa oral	Ouvinte	Apresentar interpretação pessoal de uma música em libras
41	LIBRAS - QUÃO GRANDE É O MEU DEUS / PB	Música	HINO: QUÃO GRANDE É O MEU DEUS CANTADO POR: SORAYA MORAES INTERPRETADO EM LIBRAS POR: RENATO INTERPRETAÇÃO LIBRAS EM CONTEXTO	http://www.youtube.com/watch?v=5Zd8Nmi10I	Libras e língua portuguesa oral (música) simultaneamente	Ouvinte	Apresentar interpretação pessoal de uma música em libras
42	Curso de Libras - Verbos em conflitos: Olhar, Ler e Ver - Prof. Luiz Albérico Falcão	Educação	Nossa formação destina-se às pessoas ouvintes, pais, educadores, profissionais em geral que buscam compreender a Libras e o que considerar em termos de limites, valores, conflitos, discursos e sofismas. A qualidade, variedade e a ausência de muitos sinais para a maioria dos conhecimentos que envolvem os conteúdos escolares uma panacéia	http://www.youtube.com/watch?v=4MfmV9d-Q32U	Língua portuguesa oral com alguns exemplos em libras	Ouvinte	Ensinar língua de sinais para ouvintes

43	Marisa Monte - "Depois" - Libras	Música	Musica em Libras Traduzido por Vinicius Aiache	http://www.youtube.com/watch?v=buBP9YPmYLM	Libras e língua portuguesa	Ouvinte	Apresentar interpretação pessoal de uma música em libras
44	Curso de Libras aula 17 (horas em contexto)	Educação	http://jesusestac.haman-do.blogspot.com ... Curso de Libras (basico) Deus abençoe a Todos. Visitem meu blog	http://www.youtube.com/watch?v=buBP9YPmYLM	Libras	Surdos	Exercícios para alunos ouvintes aprendizes de libras
45	CES - Piada em Libras	Educação	Piada de hoje: O toureiro.	http://www.youtube.com/watch?v=buBP9YPmYLM	Libras e LP em legenda	Surda	Entretenimento para surdos
46	LIBRAS - Vestuário Principal	Educação	Vídeo Feito para trabalho do Profº Alexandre Elias e Profª Silvia da faculdade UNIESP Sorocaba, na disciplina de Libras e Linguagens e Mediações Tecnológicas na Educação.	http://www.youtube.com/watch?v=tQBxS3a6zI4	Libras e legenda em língua portuguesa em parte do vídeo	Ouvintes aprendizes de libras	Apresentar glossário da libras como conteúdo aprendido no curso, ensinar libras

47	Fotossíntese (com Libras)	Educação	<p>Vídeo que apresenta o processo de fotossíntese.</p> <p>Este programa foi produzido pelo Projeto EMBRIO, da Universidade Estadual de Campinas com recursos do FNDE, MCT e MEC.</p>	http://www.youtube.com/watch?v=QprxVlDwgo	Língua portuguesa a com legenda em libras	Ouvintes e interpretação em libras	Apresentar conteúdo de biologia para ouvintes com interpretação para surdos
48	Há muito tempo atrás (LIBRAS)	Música	<p>Intérprete Hélio Fonseca de Araújo.</p> <p>Essa foi o Professor Valdir Balbuena que me ensinou.</p> <p>Espero que gostem.</p>	http://www.youtube.com/watch?v=TVcxKnIj0g	Libras e língua portuguesa-oral (música) simultânea-	Ouvinte	Apresentar interpretação pessoal de uma música em libras

49	Curso de Libras: Classificadores e Marcadores - Aula 1.2 - Prof. Luiz Albérico Falcão	Educação	Os classificadores e marcadores são recursos utilizados na construção do diálogo em Libras e carecem de compreensão quanto aos conceitos e valores linguísticos. A utilização destes recursos sem a devida tomada de consciência torna o processo comunicacional e cognitivo mais complexo de ser apreendido. li-bras.pernambuco@gmail.com	http://www.youtube.com/watch?v=mdEdDCZSUnk	Língua portuguesa oral com alguns exemplos em libras	Ouvinte	Ensinar língua de sinais para ouvintes
50	libras surdo 2013 SU-VAG 2013	Animais	BOM surdo libras NATAL RN SUVAG ;)	http://www.youtube.com/watch?v=...	Libras	Desenho	

51	Alexandre Elias - Pai Nosso em Libras (Língua Brasileira de Sinais)	Educação	Atuando junto a comunidade surda desde 1996 o Ms. em Comunicação e Cultura, Intérprete e Educador de surdos Alexandre Elias apresenta esta que é a oração mais conhecida do cristianismo: Pai Nosso. Que Deus abençoe suas mãos na prática da Libras!!!	http://www.youtube.com/watch?v=f_3a0alGKIA	Libras e língua portuguesaoral (música) simultaneamente	Ouvinte	Apresentar interpretação pessoal de uma música em libras
52	Bruna Karla - Que Bom que você chegou(em LIBRAS).wmv	Pessoas e blogs	INTERPRETAÇÃO EM LIBRAS REALIZADA NO DIA 21/08/2011 NO ESTÚDIO DO SISTEMA EDUCACIONAL CHAPLIN POR LÍVIA GOMES	http://www.youtube.com/watch?v=vc95naFTSnY	Libras e língua portuguesaoral (música) simultaneamente	Ouvinte	Apresentar interpretação pessoal de uma música em libras
53	Video Aula Hospedando Anjos LIBRAS	Educação	<i>Nenhuma descrição disponível.</i>	http://www.youtube.com/watch?v=vc95naFTSnY	Libras e língua portuguesaoral	Ouvinte	Apresentar interpretação pessoal de uma música em libras
54	LIBRAS - História da Educação de Surdos (ensaio da interpretação)	Educação	Preparando o primeiro trabalho de interpretação para a FENEIS/DF sobre História da Educação dos Surdos - alunos do curso básico de LIBRAS	http://www.youtube.com/watch?v=ECrBc2BX9ds	Libras e transcrição em Português	Ouvinte	Apresentar trabalho de curso de libras

55	Curso de Libras aula 11 (adverbio de tempo)	Educação	http://jesusestac.haman-do.blogspot.com Curso de Libras (basico) Deus abençoe a Todos. Visitem meu blog	http://www.youtube.com/watch?v=y9EVehLbJ3Y	Libras com legenda	Surdo	Glossário de libras para ouvintes
56	libras-gramática	Educação	<i>Nenhuma descrição disponível.</i>	http://www.youtube.com/watch?v=36ciunK4e68	Língua portuguesa com exemplos em	Ouvinte	Filmage de aula de libras ministrada para ouvintes com o objetivo de registro, divulgação e ensino da aula
57	Video de Libras "Vê os cravos nas mãos"	Música	Interpretação de Libras musica gospel	http://www.youtube.com/watch?v=...	Libras e língua por-	Ouvinte	Apresentar interpretação pessoal de uma música em libras
58	01-DIALOGO EM LIBRAS	Educação	<i>Nenhuma descrição disponível.</i>	http://www.youtube.com/watch?v=...	Libras com legenda em	?	Ensinar libras para ouvintes
59	Hino Nacional Libras	Pessoas e blogs	Hino Nacional em Libras com Legenda	http://www.youtube.com/watch?v=...	Libras e língua por-	Ouvinte	Apresentar interpretação pessoal de uma música em libras

60	Paula Fernandes - "Pra você!" LIBRAS - Natalia Romera	Música	Desculpe nos primeiros segundos do vídeo, houveram erros de edição. #esperamos que gostem. abraço sinaizado.	http://www.youtube.com/watch?v=iNmM1fcmjkg	Libras e língua portuguesa oral (música) simultânea	Ouvinte	Apresentar interpretação pessoal de uma música em libras
61	Profissões em LIBRAS	Educação	Você sabe reconhecer os sinais das profissões em LIBRAS (Língua Brasileira de Sinais)?	http://www.youtube.com/watch?v=JWPjEVSr1Z	Libras com legenda em português	Ouvintes aprendizes de libras	Apresentar conteúdo aprendido no curso, registro de atividades, ensinar libras
62	Identidade e Cumprimentos em Libras - ETEASD	Educação	Formas de se identificar e cumprimentar em Libras. Exercício ao fim do vídeo.	http://www.youtube.com/watch	Libras com legenda em português	Ouvintes aprendizes de libras	Apresentar conteúdo aprendido no curso, registro de atividades, ensinar libras

63	Em Breve: Filme "O caso LIBRAS"	Filmes e desenhos	Nos dias 24 e 25 de julho a Rimel Produções estará filmando o curta-metragem "O Caso Libras". O filme trata da história de um casal de surdos-mudos que discute dentro de um ônibus pela linguagem dos sinais. Apesar de estarem rodeados de pessoas, ninguém sabe ao certo por que o casal discute tanto. "O Caso Libras" trata de forma bem-humorada sobre o poder da linguagem de libras contando com um final inesperado que surpreende a todos. (...)	http://www.youtube.com/watch?v=toMlnfuyrl4	Língua portuguesa oral	Atores ouvintes e surdos	Trailler de divulgação de um curta metragem
64	Anjo Roupas Nova em LIBRAS by Marcos Cunha	Música	Pessoal espero que gostem , abraços de marcos cunha.	http://www.youtube.com/watch?v=toMlnfuyrl4	Libras e língua por-	Ouvinte	Apresentar interpretação pessoal de uma música em libras

65	PROFESSOR E INTÉRPRETE DE LIBRAS - PROFº OSMAR PEREIRA, METODISTA	Educação	No campus Rudge Ramos da Universidade Metodista o professor Osmar alegrou o encerramento da oficina de libras realizada no período de 6 a 10 fevereiro 2012.	http://www.youtube.com/watch?v=AHN-s5UonjI	Língua portuguesa oral e Língua de sinais	Ouvinte sinalizador	Registrar atividade do curso, divulgar o curso, ensinar libras
66	Alfabetização bilíngue - Libras e português - Jornal Visual	Notícias e política	08/01/2013 - Saiba porque a alfabetização bilíngue -- Libras e português -- é tão importante, e tão necessária. Esta reportagem, produzida pela equipe do programa Planeta, da Rede Minas, é parte do programa especial sobre os 10 anos da LIBRAS.	http://www.youtube.com/watch?v=OkgFDdyFgI	Língua de sinais e língua portuguesa oral	Ouvinte intérprete	Divulgar notícias jornalísticas para o público surdo
67	D-24 - Libras na Escola Regular: os educadores e suas estratégias	Entretenimento	Programa da disciplina D-24 - Conteúdos e Didática de Libras - do curso de Pedagogia Unesp / Univesp. Como os educadores da	http://www.youtube.com/watch?v=0YtZIV5-IQ	Língua portuguesa com filmagem de algumas	Ouvintes	Apresentar e discutir questões educacionais para surdos
68	BPC COM LIBRAS VIDEO 2 BLOCO 1	Educação	<i>Nenhuma descrição disponível.</i>	http://www.youtube.com/watch	Língua portuguesa oral	Ouvinte e intérprete de	Apresentar projeto institucional para professores surdos

69	Musica A Paz - Roupas - LIBRAS	Notícias e política	Musica A Paz - Roupas Nova, em LIBRAS, feito por Samantha Mara do Nascimento - Professora e Interprete de LIBRAS. Julho/2011.	http://www.youtube.com/watch?v=n7XepzlbU	Libras e língua portuguesa (música) simultânea	Ouvinte	Apresentar interpretação pessoal de uma música em libras
70	Prof Karina Sales- Libras- Conquistando o impossível	Educação	dia nacional do surdo!!! FACULDADE DO LITORAL SUL PAULISTA!!	http://www.youtube.com/watch?v=n7XepzlbU	Libras e língua portuguesa	Ouvinte	Apresentar interpretação pessoal de uma música em libras
71	Flor Jorge e Mateus Interpretação em Libras Andréa de Oliveira	Pessoas e blog	Interpretação em Libras da Aluna Andréa de Oliveira Conclusão do Curso de Libras do Colégio Sophus em Primavera/SP Profª Marcela Trevizan	http://www.youtube.com/watch?v=eWVSSoiXFyM	Libras e língua portuguesa simultânea	Ouvinte	Apresentar interpretação pessoal de uma música em libras

72	Língua Brasileira de Sinais Libras Aula 01 1	Educação	Graduação Pedagogia (Núcleo do Conhecimento Educação para a Diversidade) Disciplina: Língua Brasileira de Sinais: Libras Aula 01 Bloco 1 de 4 Professor(a): Silvana Elisa de Moraes Schubert / Ronaldo Quiriano da Silva Duração: <u>15:23</u> Data: 2012.10.09	http://www.youtube.com/watch?v=sMs9My5v5jk	Língua portuguesa com legenda em libras	Ouvintes	Ministrar curso de libras para ouvintes
73	Ao Mestre com Carinho em Libras	Educação	<i>Nenhuma descrição disponível.</i>	http://www.youtube.com/watch?v=...	Libras e língua portuguesa	Ouvinte	Apresentar interpretação pessoal de uma música em libras
74	Tô Passando Mal - Libras	Música	Aryta e Pedro...	http://www.youtube.com/watch?v=...	Libras e língua portuguesa	Ouvinte	Apresentar interpretação pessoal de uma música em libras
75	Aprenda Libras: Trânsito	Sem fins lucrativos/ativismo	NO QUADRO APRENDA LIBRAS, VEJA COM O INTERPRETE, VINICIUS NASCIMENTO, ALGUNS SINAIS USADOS EM UMA ESCOLA DE TRÂNSITO.	http://www.youtube.com/watch?v=FJGk2AfA30s	Língua portuguesa com legenda em libras	Ouvintes	Ensinar alguns sinais de libras em quadro de programa de TV

76	Libras UniRio	Educação	Oi! Boa tarde! Meu nome é Nathalia e eu tenho 20 anos. Eu estudo teatro na faculdade UniRio. Esse vídeo é a minha atividade de libras.	http://www.youtube.com/watch?v=g2P8sl_oBz8	Libras	Ouvinte aprendiz de libras	Apresentar atividade de libras realizada em curso
77	Pra Você - Paula Fernandes (Interpretação LIBRAS)	Entretenimento	Interpretação da música "Pra Vc" de Paula Fernandes Por Prof. Bruno Ramos da Silva	http://www.youtube.com/watch	Libras e língua portuguesa portuguesa	Ouvinte	Apresentar interpretação pessoal de uma música em libras
78	Gabriela Rocha - Nada Além de Ti (Libras)	Educação	Tive uma experiência única com Deus interpretando essa música... Sei que essas escolhas lexicais foram certas dessa vez...	http://www.youtube.com/watch?v=0YhFzQkRi8	Libras e língua portuguesa (música)	Ouvinte	Apresentar interpretação pessoal de uma música em libras
79	Libras: Alfabeto e números	Educação	Alfabeto e números em Libras	http://www.youtube.com/watch	Libras com legenda em	Ouvinte	Ensinar o alfabeto em libras

80	A IMPOR- TÂNCIA DA CO- MUNHÃO - Palestra em Libras, com áudio	Educação	Palestra em vídeo, produzida em Libras e traduzida simultaneamente para o português, em áudio. Temas básicos como batismo nas águas, Santa Ceia e comunhão com Deus e com a igreja são explicados didaticamente, com o apoio de imagens.	http://www.youtube.com/watch?v=Uvu7mefY0i4	Libras com interpretação em LP oral	Ouvinte	Ensinar conteúdo religioso para surdos
81	Curso de Libras Intermediário	Notícias e política	Curso de Libras Intermediário	http://www.youtube.com/watch?v=PMesOVL00	Libras com legenda em	Surda	Divulgar curso de libras
82	Libras: sinais do Novo Testamento.wmv	Pessoas e blogs	Este é um vídeo antigo, mas pode lhe ser útil	http://www.youtube.com/watch?v=PMesOVL00	Libras com legenda em	??	Apresentar Glossário de sinais religiosos
83	Pirâmide Alimentar em Libras	Pessoas e blogs	<i>Nenhuma descrição disponível.</i>	http://www.youtube.com/watch?v=PMesOVL00	Libras com legenda em portu-	Ouvintes aprendizes de libras	Realizar atividade do curso, divulgar informação de utilidade pública para surdos
84	Libras: livros do Antigo Testamento	Educação	sinais eclesíásticos: livros do Antigo Testamento	http://www.youtube.com/watch?v=PMesOVL00	Libras com legenda em	Ouvinte	Apresentar glossário de sinais religiosos em libras

85	Seminário de LIBRAS-UCB	Música	Vídeo sobre a importância das Instituições Públicas formarem profissionais com noções básicas em LIBRAS para buscar o tão sonhado princípio da igualdade.	http://www.youtube.com/watch?v=8e5n3fGazY	Língua portuguesa, Libras com legenda e português	Ouvintes	Realizar atividade de curso de libras para demonstrar aprendizado, divulgar problemas enfrentados
86	Video Aula - Orfãos de Deus - LIBRAS.wmv	Pessoas e blogs	<i>Nenhuma descrição disponível.</i>	http://www.youtube.com/watch?v=8e5n3fGazY	Libras e língua portuguesa	Ouvinte	Apresentar interpretação pessoal de uma música em libras
87	PROF. GABRIEL LIBRAS - O MUNDO AINDA TÃO COMPLICADO - LEGIÃO URBANA	Entretenimento	<i>Nenhuma descrição disponível.</i>	http://www.youtube.com/watch?v=zVGIjQ_Ht4U	Libras e língua portuguesa (música)	Ouvinte	Apresentar interpretação pessoal de uma música em libras
88	PROF. GABRIEL LIBRAS - AMOR I LOVE YOU - MARISA MONTE	Música	CURSO DE LIBRAS CARGA HORÁRIA - 80 HORAS (por semestre) AULAS AOS SÁBADOS	http://www.youtube.com/watch?v=ly2rnmV38TI	Libras e língua portuguesa	Ouvinte	Apresentar interpretação pessoal de uma música em libras
89	Conquistando o Impossível - LIBRAS - Marcela Trevizan.wmv	Pessoas e blogs	Aula de Libras no Colégio Sophus dia 26/10/2012 Interpretação da música Conquistando o Impossível com a Profª Marcela Trevizan	http://www.youtube.com/watch?v=BGx6bwi-rss	Libras e língua portuguesa (música) simultânea	Ouvinte	Apresentar interpretação pessoal de uma música em libras

90	Software faz a tradução da Libras para o português - Jornal Visual	Notícias e política	11/01/2013 - Para fechar a série comemorativa dos 10 anos da Língua Brasileira de Sinais, uma novidade: um software, que poderá ser instalado no telefone celular, para fazer a tradução da Libras para o português e vice versa.	http://www.youtube.com/watch?v=e65LZxZTqto	Libras e Língua portuguesa oral	Ouvinte e intérprete de libras	Divulgar notícias jornalísticas para surdos
91	PROF. GABRIEL LIBRAS - ABECE-DÁRIO DA XUXA	Música	CURSO DE LIBRAS CARGA HORÁRIA - 80 HORAS (por semestre) AULAS AOS SÁBADOS	http://www.youtube.com/watch?v=gplavq_s4fY	Libras e língua portuguesa oral	Ouvinte sinalizador	Filmagem em curso de libras interpretando música em libras

92	Curso de Libras - Expressão Facial e Corporal - Aula 2.1 - Prof. Luiz Albérico Falcão	Educação	Propomos uma metodologia vivenciada e comentada sobre o que já existe com o tema expressões faciais. Nosso foco é a formação de pessoas ouvintes em Libras, daí temos a preocupação de comentar e provocar reflexões que promovam a autonomia dos colaboradores, interlocutores e receptores diante das diversas situações de comunicação, interação, educação e instrumentalização do conhecimento humano como universal, comum e necessário para todas as pessoas: surdas e ouvintes.	http://www.youtube.com/watch?v=ImyLooTx_iU	Língua portuguesa oral com alguns exemplos em libras	Ouvinte	Ensinar língua de sinais para ouvintes
93	Como evitar gravidez LIBRAS	Educação	Esse vídeo mostra diversos métodos anti-concepcionais.	http://www.youtube.com/watch?v=7aUuwIWToY4	Libras	Ouvintes	Registrar proficiência em libras adquirida em curso, divulgar informações de saúde para surdos

94	DICIONÁRIO DE LIBRAS/POR TUGUÊS.	Animais	DICIONÁRIO DE LIBRAS / DICIONÁRIO DE PORTUGUÊS.	http://www.youtube.com/watch?v=uhLrZu5OAc	Libras	?	Informar surdos sobre questões linguísticas de um dicionário de libras e de português
95	Aula de Libras com professor Nelsão	Educação	Aula de Libras com professor Nelsão	http://www.youtube.com/watch	Língua portuguesa oral	Ouvinte	Filmagem de aula para registro, divulgação e ensino da libras
96	Carta pra Você em LIBRAS	Música	Curta nossa fan page no Facebook: https://www.facebook.com/pages/Hands-...	http://www.youtube.com/watch?v=nEXrDEsNW	Libras e língua portuguesa oral	Ouvinte	Apresentar interpretação pessoal de uma música em libras
97	Versos simples em Libras- Curso de Pedagogia PUCPR	Pessoas e blogs	Apresentação da Musica Versos Simples em Libras por Kayane e Débora.	http://www.youtube.com/watch	Libras e língua portuguesa	Ouvinte	Apresentar interpretação pessoal de uma música em libras
98	Hino Nacional Brasileiro em LIBRAS	Educação	Professor Valdir Balbueno interpretando Hino Nacional Brasileiro em LIBRAS	http://www.youtube.com/watch	Libras e língua portuguesa	Ouvinte	Apresentar interpretação pessoal de uma música em libras

99	Alfabetização bilíngue - Libras e português - Jornal Visual	Notícias e política	08/01/2013 - Saiba porque a alfabetização bilíngue -- Libras e português -- é tão importante, e tão necessária. Esta reportagem, produzida pela equipe do programa Planeta, da Rede Minas, é parte do programa especial sobre os 10 anos da LIBRAS.	http://www.youtube.com/watch?v=O-kgFDdyFgI	Libras e língua portuguesa	Ouvinte intérprete	Apresentar conteúdo de notícia/jornalístico para surdos
100	Duas metades em LIBRAS	Música	<i>Nenhuma descrição disponível.</i>	http://www.youtube.com/wat	Libras e língua por-	Ouvinte	Apresentar interpretação pessoal de uma música em libras