

UNIVERSIDADE DE SÃO PAULO
FACULDADE DE FILOSOFIA, LETRAS E CIÊNCIAS HUMANAS
DEPARTAMENTO DE LETRAS MODERNAS
PROGRAMA DE PÓS-GRADUAÇÃO EM ESTUDOS
LINGÜÍSTICOS E LITERÁRIOS EM INGLÊS

TARCÍSIO DE ARANTES LEITE

**A segmentação da língua de sinais brasileira (libras):
Um estudo lingüístico descritivo a partir da
conversação espontânea entre surdos**

São Paulo

2008

UNIVERSIDADE DE SÃO PAULO
FACULDADE DE FILOSOFIA, LETRAS E CIÊNCIAS HUMANAS
DEPARTAMENTO DE LETRAS MODERNAS
PROGRAMA DE PÓS-GRADUAÇÃO EM ESTUDOS
LINGÜÍSTICOS E LITERÁRIOS EM INGLÊS

**A segmentação da língua de sinais brasileira (libras):
Um estudo lingüístico descritivo a partir da
conversaçãõ espontânea entre surdos**

Tarcísio de Arantes Leite

Tese apresentada ao Programa de Pós-Graduação em Estudos Lingüísticos e Literários em Inglês, do Departamento de Letras Modernas da Faculdade de Filosofia, Letras e Ciências Humanas da Universidade de São Paulo, para obtenção do título de Doutor em Letras.

Orientador: Prof. Dr. Leland E. McCleary

São Paulo

2008

Sumário

1. Introdução.....	9
2. Noções básicas sobre as línguas de sinais	20
2.1. Introdução.....	20
2.2. Fonologia.....	20
2.3. Morfologia.....	26
2.4. Sintaxe	28
2.5. Prosódia	30
2.6. Gestualidade	33
2.7. Conclusão	45
3. Fundamentação teórica I: Língua e discurso na perspectiva da análise da conversa	47
3.1. Introdução.....	47
3.2. A segmentação da língua e a troca de turnos	49
3.3. A proposta original de SSJ	51
3.4. A projetabilidade dos turnos-na-fala	55
3.5. Uma proposta sobre os conceitos-chave	64
3.5.1. Turnos e TCUs.....	67
3.5.2. TRPs.....	79
3.6. Estrutura, contingência e as unidades mínimas do discurso	89
3.7. Conclusão	94
4. Fundamentação teórica II: Língua e discurso na perspectiva da gramática baseada no uso	96
4.1. Introdução.....	96
4.2. As unidades entoacionais	97
4.3. O enquadramento da atenção no discurso	106
4.4. O papel da salientação na experiência humana	112
4.5. Salientação, gestualidade e atenção conjunta.....	120
4.6. O potencial universal das unidades entoacionais	124
4.7. Conclusão	130
5. Metodologia: O desenvolvimento do corpus.....	133
5.1. Introdução.....	133
5.2. Gravação das conversas.....	138
5.3. Manipulação dos vídeos	140
5.4. Transcrição dos dados	142
5.5. Produção de relatórios	150
5.6. Conclusão	153

6.	Análise I: A troca de turnos na libras	156
6.1.	Introdução.....	156
6.2.	O caráter problemático das sobreposições de fala.....	158
6.3.	Estudos sobre conversação em línguas de sinais.....	162
6.4.	Sobreposições de fala na libras	173
6.4.1.	A segmentação da cadeia de fala sinalizada	173
6.4.2.	Circunstâncias de fala simultânea excluídas.....	182
6.4.3.	Recursos gestuais para o gerenciamento de sobreposições	185
6.5.	Conclusão	197
7.	Análise II: As unidades entoacionais na libras.....	203
7.1.	Introdução.....	203
7.2.	Procedimentos de análise	204
7.3.	Instâncias de turnos simples	211
7.3.1.	Alongamento final	212
7.3.2.	Aceleração inicial.....	216
7.3.3.	Gestos atencionais coesos	220
7.3.4.	Piscada de olhos.....	221
7.3.5.	Acento enfático	222
7.4.	Instâncias de turnos complexos estruturados	223
7.4.1.	Listas	223
7.4.2.	Contrastes.....	237
7.5.	Conclusão	252
8.	Conclusão	258
9.	Bibliografia.....	265

DEDICATÓRIA

Esta tese é dedicada aos surdos e à sua luta
pelo reconhecimento pleno das línguas de sinais

AGRADECIMENTOS

- Ao meu mentor intelectual, prof. Leland McCleary, por me mostrar que sempre é possível ir além;
- À minha professora Evani Viotti, pela orientação informal ao longo da minha formação;
- Aos professores José Luiz Fiorin, Sandra Madureira, Cláudia de Lemos, Ronice Quadros e Pedro Garcez pelas críticas e sugestões nas discussões da qualificação e/ou da defesa;
- A todos os colegas do grupo de pesquisa da USP, antropólogos, linguistas e historiadores, por ajudarem a ampliar meus horizontes teóricos;
- À Coordenação de Aperfeiçoamento de Pessoal de Nível Superior (CAPES), pelo financiamento desta pesquisa;
- Ao Conselho Nacional de Desenvolvimento Científico e Tecnológico (CNPq), pelo financiamento do projeto de formação de corpus da libras que correu paralelamente ao desenvolvimento de minha tese;
- À Escola do Futuro/USP, por disponibilizar um espaço de trabalho ao nosso grupo de pesquisa;
- A Han Sloetjes e toda a equipe de elaboração do ELAN, pelo trabalho sem o qual essa pesquisa não seria viável;
- A Regiane P. Agrella, Wilson S. Silva, Sandro dos S. Pereira, por me cederem gentilmente o seu tempo na pesquisa e compartilharem comigo o seu conhecimento da libras;
- A Cristiano de Castro A. Koyama e Andrea Iguma, pelas dicas na produção dos vídeos lincados à tese;
- A Pedro Augusto Marques por gentilmente ceder as imagens do dicionário libras on-line, e a João Luis Nascimento, sinalizador do dicionário;
- A toda a equipe do Centro de Computação Eletrônica (CCE) da USP, por viabilizarem estúdio, iluminação e câmeras para a gravação das conversações;
- A minha mãe, pai e irmãos, pelo amor e carinho sempre;
- A Cris, pelo amor e companheirismo nesses últimos quatro anos;
- E aos amigos eternos, pela brodagem que o tempo só fortalece.

RESUMO

A presente tese consiste numa pesquisa de caráter lingüístico descritivo. O objetivo é o de oferecer critérios para a segmentação do discurso na língua de sinais brasileira (libras) em unidades gramaticais. Duas linhas teóricas contribuíram para este projeto de forma crucial: a análise da conversa de base etnometodológica e a gramática baseada no uso. A análise da conversa, ao observar o modo como os próprios participantes se orientam uns em relação aos outros na interação, revela uma visão dos recursos gramaticais como práticas sociais voltadas à realização e coordenação de ações sociais na conversação. A gramática baseada no uso, ao explorar as relações entre língua, cognição e interação, revela o importante papel da prosódia e da gestualidade para o estabelecimento da atenção conjunta na interação. Na constituição do corpus, díades formadas por surdos adultos fluentes em libras foram gravadas num estúdio, com iluminação e câmeras posicionadas de modo a captar o rosto e o espaço de sinalização dos falantes, além de ambos em perfil. Um trecho da gravação foi selecionado para ser transcrito e utilizado como ponto de referência principal da análise, que envolveu duas etapas distintas. Na primeira fase, é demonstrado que o princípio do *um-de-cada-vez*, uma manifestação da atenção conjunta no nível do discurso, também se mostra operante numa língua de modalidade gestual-visual como a libras. Na segunda fase, uma análise sobre a segmentação interna dos turnos de fala sinalizados é realizada por meio da combinação de uma abordagem com foco nas unidades entoacionais do discurso, e uma abordagem com foco em práticas estruturadas do discurso, tais como listas e contrastes. As principais contribuições que resultaram do trabalho foram: i) um repertório sistematizado de recursos manuais e não-manuais da libras que pode servir como ponto de referência inicial para a segmentação do discurso espontâneo em unidades gramaticais; e ii) uma abordagem para a segmentação do discurso que, em futuras investigações, poderá permitir um maior aprofundamento de nosso conhecimento acerca dos recursos prosódicos da libras.

Palavras-chave: segmentação; conversação; atenção conjunta; unidade entoacional; língua de sinais brasileira.

ABSTRACT

This dissertation comprises a study in descriptive linguistics. The aim is to elaborate criteria for the segmentation of Brazilian Sign Language (Libras) discourse into grammatical units. Two theoretical fields contributed crucially to this project: conversation analysis and usage-based grammar. Conversation analysis, by focusing observation on how participants themselves orient to each other in interaction, has revealed a perspective on grammatical resources as social practices designed for the deployment and coordination of social actions in conversation. Usage-based grammar, by exploring the relations among language, cognition and interaction, reveals the important role of prosody and gesture for the establishment of joint attention in interaction. In the construction of the corpus, dyads composed of fluent adult speakers of Libras were recorded in a studio, with lights and cameras positioned in order to capture both the face and the signing space of each of the speakers, in addition to the two speakers in profile. An excerpt from one of the conversations was selected to be transcribed and used as the main point of reference for the analysis, which involved two distinct phases. In the first phase, it is demonstrated that the principle of *one-at-a-time* – a manifestation of joint attention at the level of discourse – also holds in a language of gestural-visual modality like Libras. In the second phase, the internal segmentation of turns at talk is analyzed, by combining an approach focused on the minimal prosodic units of discourse with an approach focused on structured practices of discourse, such as lists and contrasts. The major contributions resulting from the analysis are: i) a systematized repertoire of manual and non-manual resources of Libras which can serve as an initial point of reference for the segmentation of signed discourse into grammatical units; and, ii) an approach to the identification of prosodic units which, in future investigations, may permit further advancements with respect to our understanding of the prosodic resources of Libras.

Key-words: segmentation; conversation; joint attention; intonation unit; Brazilian Sign Language.

“Quando nada acontece,
há um milagre que não estamos vendo”

João Guimarães Rosa

1. Introdução

Conversar é uma das atividades mais básicas do ser humano. É tão básica que se constitui num pré-requisito para a realização de inúmeras outras atividades importantes para a experiência humana. Desde as tarefas mais simples, como comprar um suco na padaria da esquina, até as mais complexas, como governar um país de dimensões continentais como o Brasil, encontros conversacionais com diferentes níveis de complexidade se mostram fundamentais tanto para o sucesso quanto para o fracasso dos mais diversos empreendimentos sociais humanos.

Apesar disso, é comum a visão da conversação como uma atividade nula, um meio fútil de “fazer hora” quando não há nada melhor para se envolver. Eu poderia apostar, por exemplo, que o leitor já deparou com uma situação similar à que se segue. Estão dois amigos sentados em algum lugar, conversando, quando outro amigo passa coincidentemente pelo local e pergunta: “E aí, tudo bem? O que estão fazendo?”. Embora várias coisas possam, ou devam estar sendo feitas, uma resposta bastante comum de se ouvir é: “*Nada*, estamos só conversando”.

Se “conversar” pode ser entendido como “não fazer nada”, talvez devamos fazer valer a sabedoria da epígrafe deste texto, buscando entender o “nada” não como uma ausência de fatos notáveis, mas como uma falta de sensibilidade, ou capacidade, ou perspicácia de notarmos o que de mais importante acontece em nossa vida cotidiana. Quando estudamos com o devido cuidado e rigor, por exemplo, como é que duas pessoas procedem para realizar uma conversa das mais corriqueiras, chama-nos a atenção: a enorme variedade dos recursos empregados; a complexidade na estruturação desses recursos para a organização da participação; o tato com o qual ações sociais são realizadas; a flexibilidade no modo de lidar com as contingências; o alinhamento temporal fino das participações; a grandeza do conhecimento de mundo que é implicitamente evocado para produzir e compreender a fala, entre outros vários aspectos. Tal análise revela, portanto, que, quando “não estamos fazendo nada, só conversando”, estamos na verdade engajados numa complexa atividade colaborativa que requer um grande investimento em termos de conhecimento, atenção e interesse por

todos aqueles que se ratificam mutuamente como participantes de um encontro conversacional.

A presente tese traz como tema a abordagem de uma das várias facetas dessa complexa tarefa: a questão da segmentação da conversação em unidades gramaticais, com foco específico sobre a língua de sinais brasileira (libras). Tal questão não é apenas um construto do mundo teórico do analista; trata-se de um problema prático dos participantes de uma interação. Identificar e interpretar as unidades de uma língua é fundamental, em primeiro lugar, para compreender as ações que os falantes realizam por meio dessa ferramenta social que chamamos de “língua”, e, em segundo lugar, para coordenar temporalmente as ações de um interlocutor com as de outro, de modo a possibilitar uma interação eficaz. Para os próprios participantes de uma conversa, porém, os recursos formais que realizam essa segmentação são produzidos e interpretados inconscientemente, de modo que ao pesquisador cabe a tarefa de identificá-los, explicitando o modo como eles interagem em contextos específicos de fala a fim de permitir a projeção de cursos subseqüentes de ação.

No que se refere à relevância mais especificamente lingüística dessa análise, a existência de critérios robustos de segmentação do discurso se mostra fundamental para a descrição das línguas naturais. Por razões que serão discutidas em maior profundidade no capítulo 3 e principalmente no 4, abordagens lingüísticas que compartilham uma visão de gramática baseada no uso têm privilegiado a identificação da organização prosódica da fala como caminho fundamental para chegar às unidades gramaticais de uma língua. No que diz respeito às línguas orais (LOs), essas unidades entoacionais têm sido estudadas extensamente nas últimas décadas, a ponto de pesquisadores treinados serem capazes de identificar agrupamentos prosódicos mínimos até mesmo em línguas por eles desconhecidas, isto é, sem um conhecimento específico sobre o conteúdo do que está sendo dito.

Um novo desafio surge, então, quando os lingüistas se dão conta de que as línguas de sinais (LSs), realizadas por meio de um canal gestual-visual, são também línguas naturais. Na década de 60, o primeiro estudo científico de uma língua de sinais – a língua de sinais americana (ASL) – fez uso dos mesmos princípios de análise estruturalistas, que se mostravam hegemônicos na época, para demonstrar que as LSs compartilhavam das mesmas propriedades universais das línguas naturais, tais como a produtividade e a arbitrariedade. Nas décadas subseqüentes, estudos voltados para as

LSs permaneceram seguindo essa linha básica de raciocínio: se as LSs são línguas naturais, elas devem apresentar os mesmos mecanismos de funcionamento que as LOs apresentam.

Seguindo essa mesma tendência, poderíamos assumir que as unidades gramaticais nas LSs fossem formadas por sintagmas, orações e sentenças com características estruturais semelhantes às das LOs. Alguns pesquisadores, contudo, têm sugerido cautela a esse respeito. Charlotte Baker e Carol Padden, já há quase três décadas, alertavam sobre o risco de assumir o conhecimento da estrutura das LOs para a análise das LSs:

We do not yet have a grammar of ASL. One major obstacle toward our devising such a grammar is that we do not really know what to call a sentence in ASL or what constitutes a grammatical unit. This is particularly true of those signed sequences that involve mechanisms not found in oral languages – such as the signer’s use of space to indicate grammatical relationships. We want to be wary of imposing English or oral language-based structure on the sign. What we need is a tool for understanding what deaf signers perceive to be grammatical units in their language (Baker e Padden, 1978: 35).

Apesar desse alerta, o questionamento sobre a natureza das unidades gramaticais nas LSs tem sido abordado com pouca frequência na literatura.¹ De maneira geral, predominam no campo metodologias clássicas, tais como a eliciação de sentenças (i.e. solicitação ao informante bilíngüe que ofereça traduções de sentenças-alvo construídas) e o apelo à intuição de falantes (i.e. recurso aos julgamentos de gramaticalidade diante de sentenças também construídas). Em ambos os casos, o pesquisador assume que as sentenças traduzidas ou julgadas constituam unidades gramaticais válidas ou próprias da língua de sinais estudada.

Embora tais metodologias tenham um papel importante no complexo trabalho de investigação gramatical das línguas naturais, a questão problemática sobre o seu uso em pesquisas com LSs se refere à sua *primazia* em relação à análise de dados espontâneos, considerando-se o viés em potencial que as LOs (e principalmente a escrita) podem imprimir sobre os dados sinalizados. A adoção de tais metodologias como centrais nos

¹ Para uma exceção, ver o workshop recentemente realizado na Alemanha, em 2006, cujos trabalhos podem ser encontrados no seguinte endereço (ainda que apenas na forma resumida): <http://www.let.ru.nl/sign-lang/sentence/>.

trabalhos acadêmicos da área, ou ainda a falta de critérios para a segmentação gramatical do discurso em trabalhos que utilizam dados espontâneos ou quase-espontâneos, envolvem o risco de tornar os resultados finais enviesados e/ou duvidosos. É em face desse contexto, então, que surge a presente proposta de pesquisa: uma busca por pistas de estruturação prosódica do discurso na libras, a partir da análise da conversa espontânea entre surdos adultos fluentes, a fim de chegar às unidades gramaticais próprias dessa língua.

Um outro modo de apresentar essa proposta é dizendo que, se pesquisas sobre a libras e sobre outras LSs têm tradicionalmente assumido os limites das unidades gramaticais a fim de que a investigação possa passar ao que “realmente interessa” (i.e. análises sobre processos fonológicos e morfológicos e operações sintáticas), o objetivo desta pesquisa é o de dar um passo para trás nessa tendência, definindo o que “realmente interessa” como uma problematização da identificabilidade de unidades gramaticais na libras antes de qualquer outra coisa. A esperança é a de que essa reflexão possa fundamentar as análises futuras a respeito das operações e processos lingüísticos aos quais tais unidades se vêm submetidas de uma maneira mais sólida.

Embora a necessidade de abertura para a singularidade das LSs possa parecer um tanto óbvia, o fato de tais línguas terem sido desprezadas, discriminadas e reprimidas por séculos – uma situação que ainda não foi completamente superada – acabou inevitavelmente colocando os pesquisadores da área numa posição de defesa: a de precisar comprovar, continuamente, que tudo o que já foi demonstrado cientificamente sobre as LOs pode também ser aplicado às LSs. Nesse sentido, a maior abertura que alguns pesquisadores têm demonstrado, mais recentemente, revela não apenas uma maior sensibilidade à empiria, produto de algumas décadas de estudo na área, mas também um certo desprendimento científico em relação às questões políticas que envolvem a área da surdez, o que me parece fundamental para o avanço da área e da ciência como um todo.

Esse avanço já começa a ser percebido no campo. A postura de “diálogo” mais franco com a empiria tem levantado uma série de questões a respeito das LSs que, quando acompanhada de uma postura reflexiva do pesquisador, exige dele uma reconsideração de seu próprio entendimento das LOs. Como resultado, muitos dos aspectos a princípio tomados como específicos das LSs acabam por revelar dimensões de análise que haviam sido ignoradas nos estudos lingüísticos, ora por razões inevitáveis

(e.g. limitação tecnológica para registro e manipulação de dados), ora por motivos de outra ordem (e.g. vieses metodológicos e teóricos).

No decorrer da leitura, é importante ter-se em mente que o objetivo central do estudo é o de contribuir para a descrição lingüística da libras. Por isso, ainda que conceitos e *insights* de áreas tais como a análise da conversa e a lingüística cognitiva sejam cruciais para a fundamentação teórica do trabalho, meu objetivo não é o de realizar uma análise de cunho sociológico ou cognitivista. No meu entendimento, a fundamentação teórica aqui trazida tem uma função específica, quase metodológica: a de oferecer ao pesquisador critérios na abordagem do discurso espontâneo em libras que permitam, a despeito de sua complexidade, viabilizar a identificação de unidades entoacionais como um primeiro passo para a identificação de unidades gramaticais.

O capítulo 2 traz um panorama sobre as pesquisas no campo das LSs, abordando brevemente alguns achados fundamentais no que tange a diferentes níveis de análise: a fonologia, a morfologia, a sintaxe, a prosódia e a gestualidade. Trata-se de um capítulo importante principalmente para aqueles que ainda têm pouco contato com a gramática dessas línguas, embora a sua leitura também deva ser aproveitada por aqueles que, mesmo estando familiarizados com essa literatura, desconhecem a nova perspectiva que vem sendo construída por pesquisadores tais como Scott Liddell e Sherman Wilcox, entre outros: um olhar sobre a gramática das LSs que revela o papel fundamental da gestualidade. O objetivo geral desse panorama, então, é o de oferecer uma base de apoio importante para a leitura dos capítulos de análise, principalmente o último, que trata mais especificamente da relação íntima entre gestualidade e prosódia na libras.

Os capítulos 3 e 4 estão intimamente relacionados, mas acabaram divididos em dois por motivos didáticos. Ambos tratam da perspectiva sobre língua e discurso que é adotada nesta tese, a primeira parte revelando a contribuição mais específica da análise da conversa e da lingüística interacional para essa perspectiva, e a segunda parte, a da lingüística funcional e cognitiva e dos estudos da gestualidade. Como ficará claro na leitura desses capítulos, quando uma das abordagens estiver sob discussão, elementos da outra serão trazidos aqui e ali a fim de oferecer uma perspectiva crítica. A proposta é a de iniciar uma articulação entre elas, apontando de que maneira pontos de fraqueza em uma podem ser suplantados pelo olhar da outra. Tal articulação é possível, a meu ver, devido a uma complementaridade nas explicações que cada uma dessas abordagens oferece a um problema central para a questão da segmentação do discurso: como é que

somos capazes de coordenar nossa atenção na interação a um mesmo “objeto” a cada novo momento, de modo a tornar essa interação bem-sucedida.

No capítulo 3, a resposta a essa questão envolve uma discussão sobre a construção das unidades mínimas do discurso e a projetabilidade de seus possíveis pontos de completude, aspecto responsável pela organização da conversação sob o princípio do “um-de-cada-vez”. A discussão recupera, inicialmente, a proposta original do sociólogo Harvey Sacks e seus colaboradores sobre o conceito de *unidade de construção de turnos* como recurso primordial frente ao qual os participantes orientam a sua participação na conversação. Em seguida, traz a contribuição de alguns pesquisadores, a maioria deles do campo recente da lingüística interacional, que seguiram a trilha aberta por Sacks e colegas e aprofundaram a nossa compreensão acerca do fenômeno da projetabilidade, refletindo sobre a intrincada relação entre recursos formais diversos, de um lado, e seqüências conversacionais específicas, de outro, no gerenciamento da troca de turnos.

A proposta da presente pesquisa diante dessa literatura é apresentada em seguida. Argumenta-se a favor de uma visão sobre as unidades da fala menos enviesada pela escrita, salientando-se a importância da prosódia e da gestualidade como recursos fundamentais para a construção e o gerenciamento da participação na interação e, conseqüentemente, para a delimitação de unidades gramaticais. Além disso, propõe-se uma visão mais flexível sobre conceitos-chave da teoria, tais como “turnos”, “unidades de construção de turnos” e “ponto de relevância para transição entre falantes”. Essa flexibilidade decorre do fato de uma fala situada poder envolver, dependendo dos recursos formais empregados e da seqüência de fala imediata, diferentes graus de participação na interação e diferentes graus de relevância para transição entre falantes em pontos discretos do turno, o que vem ressaltar tanto a dimensão estrutural quanto a dimensão contingencial do trabalho de projeção da fala.

O capítulo 4 vem complementar a discussão do capítulo 3, aprofundando a reflexão sobre a importância da prosódia e da gestualidade para a coordenação de ações sociais na conversação, bem como sobre a correlação entre prosódia e gramática. Inicialmente, a discussão recupera o conhecimento acumulado nos campos da lingüística funcional e cognitiva sobre as características formais e funcionais das unidades entoacionais. Tal discussão permite, então, situar a função dessas unidades num âmbito maior, relacionado ao fenômeno da salientação na experiência humana: o

estabelecimento de relações de figura-e-fundo, com a emergência de uma “entidade” como *foco* da atenção enquanto outras são relegadas ao *fundo* da atenção. Serão, então, trazidas evidências de que esse fenômeno permeia diferentes âmbitos da experiência humana, tal como a percepção, a cognição, a língua e a interação – o princípio do um-de-cada-vez da conversação sendo apenas uma dessas manifestações, no âmbito da interação.

Além disso, serão também trazidas evidências de que, na conversação, o estabelecimento da atenção conjunta sobre um único “objeto” a cada momento é um passo fundamental para a aquisição das línguas naturais, como têm revelado os estudos de Michael Tomasello e colaboradores. Nessa discussão, ficará mais evidente a relação íntima entre gestualidade e prosódia, com ambos os domínios envolvendo a exploração de gestos articulatórios – ora concomitantes com a fala, ora seqüenciados – empregados a fim de direcionar a atenção do interlocutor a uma dada entidade em detrimento de outras. O capítulo se encerra, então, com uma discussão sobre o potencial universal das unidades entoacionais, agrupando os achados de pesquisas realizadas com diversas LOs historicamente não-relacionadas e reforçando, dessa maneira, a pertinência da utilização desse conceito como ponto de apoio importante para a análise de uma língua como a libras.

Terminados os dois capítulos de articulação teórica da tese, a discussão passa, no capítulo 5, a tratar dos aspectos metodológicos da pesquisa. A reflexão aborda, inicialmente, a questão dos diferentes tipos de dados que podem ser obtidos na pesquisa lingüística a partir da utilização de procedimentos metodológicos específicos. Argumenta-se – *principalmente* para o caso das LSs, embora não *exclusivamente* – a favor da importância de utilização de dados naturais como fonte primária de análise, por meio de gravação de falantes em situações espontâneas ou quase-espontâneas de uso da língua. Esse favorecimento emerge num momento histórico que possibilita a exploração desse tipo de metodologia, com o surgimento de tecnologias de registro e manipulação de dados cada vez mais variadas e eficientes, tornando pesquisas pautadas no discurso espontâneo amplamente viáveis.

Em seguida, a discussão trata das diferentes etapas de formação do corpus (i.e. gravação, manipulação dos vídeos e transcrição), e da construção de relatórios a partir das análises. Serão apontados os procedimentos adotados em cada uma dessas etapas, os problemas que acabaram surgindo ao longo do processo, e as soluções que foram

adotadas a fim de saná-los. Tendo em vista o caráter inusitado da atual proposta de estudo (i.e. uma análise gramatical de uma língua de sinais com base na conversação), vários dos procedimentos aqui discutidos não envolvem escolhas do pesquisador diante de propostas correntes na literatura, como costuma ser o caso, mas opções criadas e testadas no próprio processo de condução da pesquisa.

Os capítulos 6 e 7, por fim, tratam da análise propriamente dita dos dados coletados. O primeiro deles, o capítulo 6, envolve uma reflexão sobre a pertinência de estudar uma língua de sinais sob a perspectiva da análise da conversa. O intuito é o de demonstrar que o princípio do um-de-cada-vez – que orienta a proposta teórica de Harvey Sacks e colaboradores e que é tomado como pressuposto central desta tese – se mostra operante também na conversa em libras. Tal demonstração é feita por meio de uma análise qualitativa sobre as repercussões sentidas nas falas dos participantes nos contextos em que esse princípio é violado, isto é, nos momentos em que mais do que um fala ao mesmo tempo.

Inicialmente, a discussão envolve uma reflexão sobre como definir sobreposições de fala problemáticas para os participantes, tendo em vista que nem todas as falas simultâneas na conversação apresentam esse caráter. Em seguida, alguns estudos voltados para a conversação no campo das LSs serão resenhados sob o olhar crítico da perspectiva teórica adotada nesta tese – isto é, considerando-se a importância do estabelecimento do um-de-cada-vez a fim de que interação seja bem-sucedida. Por fim, com a análise dos contextos da conversa em libras nos quais o princípio do um-de-cada-vez é violado, veremos que repercussões na fala dos participantes, na forma de disfluências e perturbações de vários tipos, puderam ser encontradas. A análise dessas ocorrências, então, irá sugerir que tais modificações da fala constituem recursos metódicos empregados pelos participantes para a resolução do que é tido como um problema interacional que necessita ser localmente gerenciado.

Nesse primeiro capítulo de análise, a questão da segmentação aparece mais como pano de fundo para a demonstração da relevância do princípio do um-de-cada-vez, sendo discutida principalmente no que tange à identificação de fronteiras nos turnos de fala dos participantes surdos. No capítulo seguinte, então, a questão da segmentação torna-se o foco central da discussão, com um aprofundamento na reflexão sobre a organização *interna* dos turnos sinalizados, seja em unidades entoacionais, seja em construções recursivas de maior complexidade.

A discussão do capítulo 7 se inicia com uma explicitação dos procedimentos de análise utilizados nesta etapa. Tendo em vista que o discurso espontâneo apresenta grande complexidade, critérios foram adotados para selecionar trechos específicos que pudessem facilitar a identificação das unidades mínimas do discurso, sem que isso implicasse uma abstenção de dados naturais ou mesmo uma desconsideração da dimensão interativa dos enunciados. Tais procedimentos envolveram a combinação de duas abordagens: uma sobre as unidades mínimas do discurso, aproveitando-se do conhecimento sobre a forma e função das unidades entoacionais; a outra sobre práticas discursivas estruturadas, aproveitando-se do conhecimento sobre listas e contrastes.

No que concerne à primeira abordagem, instâncias de turnos simples, possivelmente compostos por uma única unidade, foram selecionadas e analisadas assumindo-se que as características dessas instâncias, em especial no que concerne às suas fronteiras, refletiriam características de unidades entoacionais na libras. O conhecimento acumulado nessa primeira etapa, então, serviu como ponto de apoio para a análise de cadeias de fala mais longas, em que práticas sociais tais como listas e contrastes eram produzidas pelos participantes. O olhar analítico dessa segunda etapa foi então complementado pelo conhecimento acumulado no campo da análise da conversa sobre a estruturação dessas práticas em LOs, o que contribuiu para a identificação tanto de sua organização interna quanto externa (i.e. relativa ao contexto de fala imediato onde tais práticas se inserem). Nessa discussão, evidências sobre a relevância dos recursos formais identificados foram buscadas no modo como os próprios participantes reagiam diante da produção das listas e contrastes.

Como um todo, o processo de análise reportado no capítulo 7 demonstrou que achados alcançados num primeiro momento de investigação puderam ser, em momentos subseqüentes, ora corroborados, ora submetidos a uma maior generalização, a fim de dar conta dos novos dados. O resultado foi o levantamento de um repertório de recursos manuais e não-manuais que dizem respeito à estruturação prosódica na libras tanto no nível do discurso (i.e. características de unidades entoacionais), quanto no nível de suas unidades básicas (i.e. características do acento).

A conclusão encerra a tese discutindo, entre outras coisas, as principais contribuições que a pesquisa oferece. No âmbito teórico, destaca-se a proposta de articulação teórica entre os campos da análise da conversa e da linguística interacional, de um lado, e da linguística cognitiva e funcional, de outro. No que tange ao primeiro

campo, uma visão de língua e discurso menos enviesada pela escrita é ressaltada, buscando trazer o papel da corporalidade para o centro da reflexão sobre o gerenciamento da participação na interação. No que tange ao segundo campo, uma visão de língua e discurso menos mentalista é ressaltada, o que implica uma compreensão da situação de uso da língua como uma arena de ação social, envolvendo *também* a troca de informações mas se estendendo para além dela.

No âmbito metodológico, destaca-se o conhecimento produzido acerca do processo de formação de um corpus em língua de sinais. Esse conhecimento se mostra particularmente novo por envolver um gênero de fala como a conversação, o que exige do pesquisador: i) a disponibilização de um aparato tecnológico relativamente complexo para gravação de dados com boa qualidade (estúdio, câmeras, iluminação); ii) o domínio de programas de manipulação de arquivos de vídeo, seja para a sua operacionalização em computadores, seja para a melhoria da qualidade das gravações, seja ainda para a geração de relatórios que explorem as tecnologias existentes de modo a permitir uma exposição das análises da libras capaz de ser refutada; e iii) o uso de uma ferramenta de transcrição de dados multimodais capaz de captar a temporalidade da fala-em-interação e de manipular mais do que um vídeo simultaneamente.

Além disso, o procedimento de análise adotado – combinando uma abordagem com foco nas unidades entoacionais com uma abordagem com foco em práticas discursivas complexas – se mostrou profícuo para o propósito da tese, de levantamento de pistas de segmentação prosódica do discurso na libras. Isso demonstra a pertinência da proposta, que pode continuar a ser desenvolvida por meio de uma ampliação do corpus e/ou a consideração de novas práticas estruturadas, possibilitando a expansão e o aprofundamento de nosso conhecimento atual sobre os recursos prosódicos da libras.

No âmbito empírico, a tese se coloca como um dos primeiros passos neste empreendimento recente que é a descrição da libras com base numa perspectiva de gramática baseada no uso. Tendo em vista a importância da prosódia na análise da segmentação do discurso, os achados desta tese devem se mostrar relevantes, como ponto de partida, para investigações futuras sobre diferentes níveis de análise da libras: o fonético-fonológico (e.g. modulações da fala e processos fonológicos); o prosódico (e.g. as características das unidades entoacionais e do acento); e o discursivo (e.g. recursos de coesão textual, a estruturação de práticas sociais complexas, o gerenciamento da troca de turnos).

Cabe destacar, por fim, que a presente tese integra um projeto maior, de formação de corpus da libras, que vem sendo desenvolvido pelo grupo *Estudos da Comunidade Surda: Língua, Cultura, História*, coordenado pelo professor Leland McCleary. O intuito principal dos pesquisadores da área de língua, nesse projeto, é o de viabilizar uma descrição da libras baseada em corpus de natureza espontânea ou quase-espontânea. Tal trabalho deve trazer importantes contribuições teóricas e aplicadas: para o estudo lingüístico de maneira geral, nos impelindo a repensar visões já bastante consolidadas sobre as propriedades e o modo de funcionamento das línguas naturais; e, para a vida da comunidade surda, produzindo um conhecimento que será importante para a estruturação de uma educação bilíngüe, com ênfase no ensino da libras como primeira língua para surdos e como segunda língua para ouvintes.

2. Noções básicas sobre as línguas de sinais

2.1. Introdução

O objetivo deste capítulo é o de oferecer, principalmente ao leitor não familiarizado com os estudos lingüísticos das LSs, um panorama dos achados fundamentais a respeito dessas línguas que permitiram aos pesquisadores enquadrá-las na categoria de línguas naturais, juntamente com as LOs. Além disso, uma maior ênfase será dada a aspectos das LSs que devem se mostrar particularmente relevantes para as análises deste estudo, tais como a estruturação interna dos sinais manuais e as suas modulações, as formas e funções dos sinais não-manuais, e a relação entre língua de sinais e gestualidade.

A língua de sinais americana (ASL) é hoje, sem dúvida, a língua de sinais mais estudada nesse campo ainda recente, em grande parte pelo pioneirismo do trabalho de Stokoe (1960). Por esse motivo, as principais referências aqui trazidas serão relativas aos estudos da ASL, e exemplos da libras deverão ser eventualmente trazidos a título de ilustração de algumas das propostas apresentadas.² A perspectiva adotada nesta resenha é, em essência, aquela apresentada por Liddell (2003a), em especial pelo fato de esse pesquisador ter trazido um olhar renovador sobre a investigação da gramática das LSs, aproximando-a dos estudos da gestualidade.³

2.2. Fonologia

Uma das características mais fundamentais das línguas naturais é o que Martinet (1978) chamou de “dupla articulação”, isto é, a propriedade das línguas humanas de se estruturarem num plano de conteúdo (a 1ª articulação), formado por morfemas e palavras, a partir da estruturação de um segundo plano, de elementos sem conteúdo e de caráter distintivo (a 2ª articulação), formado por fonemas. Essa característica, que

² Os exemplos trazidos não devem ser tomados como evidência *prima facie* de que essas propostas se aplicam tal qual na libras. Somente um trabalho empírico abrangente, com base em um corpus robusto, poderá mostrar em que medida a libras se aproxima e em que medida se distancia dos achados propostos para a ASL e outras LSs. Isso dito, entendo que a inclusão de exemplos na libras seja válida no intuito de mostrar a relevância em potencial dessas propostas para a análise dessa língua.

³ Liddell foi inicialmente um gerativista, e alguns de seus trabalhos – inclusive alguns aqui citados – foram produzidos pelo autor sob a orientação desse modelo teórico. Sua mudança de orientação, iniciada a partir da metade da década de 80, foi motivada, segundo o autor, pela impossibilidade de esse modelo dar conta de importantes fenômenos empíricos observados na ASL (Liddell, 2003a). Alguns desses fenômenos serão discutidos na seção 2.6.

confere produtividade às línguas humanas por permitir que um número restrito de elementos seja capaz de representar um número irrestrito de significados, foi demonstrada como presente também nas LSs a partir dos trabalhos de Stokoe (1960) e Stokoe et al. (1965).⁴

Stokoe e colegas mostraram que, aplicando-se a metodologia estruturalista clássica de contraste entre pares mínimos, era possível identificar uma série de aspectos dos sinais que, em si, não possuíam significado, mas que, quando agrupados de diferentes maneiras, permitiam a distinção de itens lexicais. Stokoe chamou esses elementos de *queremas* (do grego *khéir*, mãos) por acreditar que a denominação *fonema*, com sua etimologia relacionada ao som, fosse por demais enviesada pelos estudos das LOs. Inicialmente, foram três os aspectos propostos como componentes da estrutura interna dos sinais (Stokoe, 1960): a configuração de mão, ou CM (originalmente denominada *designator*); a localização, ou L (originalmente denominada *tabula*); e o movimento, ou M (originalmente denominado *signation*).⁵

Alguns exemplos da libras podem servir como ilustração dessa proposta de Stokoe (ver figura 1 abaixo). Tomando como pares mínimos o sinal <[ACOSTUMAR e EDUCAÇÃO](#)>,^{6,7} notamos que a única diferença entre os dois está na CM, que no primeiro sinal é realizada com a configuração “B”, ao passo que no segundo é realizada com a configuração “L”.⁸ Outro par mínimo que exemplifica a análise de Stokoe é <[ANTES e ONTEM](#)>, cuja única diferença reside no parâmetro L, no primeiro caso realizado na palma da mão passiva e no segundo na bochecha. Por fim, um terceiro exemplo é o par <[SEMANA e IR](#)>, cuja única diferença reside no parâmetro M, no primeiro caso, realizado com uma trajetória retilínea e, no segundo caso, em forma de arco.

⁴ No Brasil, os estudos de LSs tiveram início a partir da década de 80, com os trabalhos de Lucinda Ferreira-Brito (e.g. 1984; 1995; ver também Revista *Geles*, organizada pela autora).

⁵ Essa terminologia original, à qual Stokoe se referia por meio das formas abreviadas “des”, “tab” e “sig”, não foi adotada por outros pesquisadores.

⁶ Tendo em vista que a libras não possui uma escrita própria, os sinais são aqui apresentados na forma de glosas, representadas em caixa alta para diferenciá-los do português.

⁷ Ao longo do texto, trechos destacados no texto por meio da marcação ‘<__>’ referem-se a vídeos linkados à versão digital da tese, contendo sinalizações de sinais isolados, exemplos e excertos discutidos.

⁸ Configuração de mão em “B”:  e configuração de mão em “L”: .

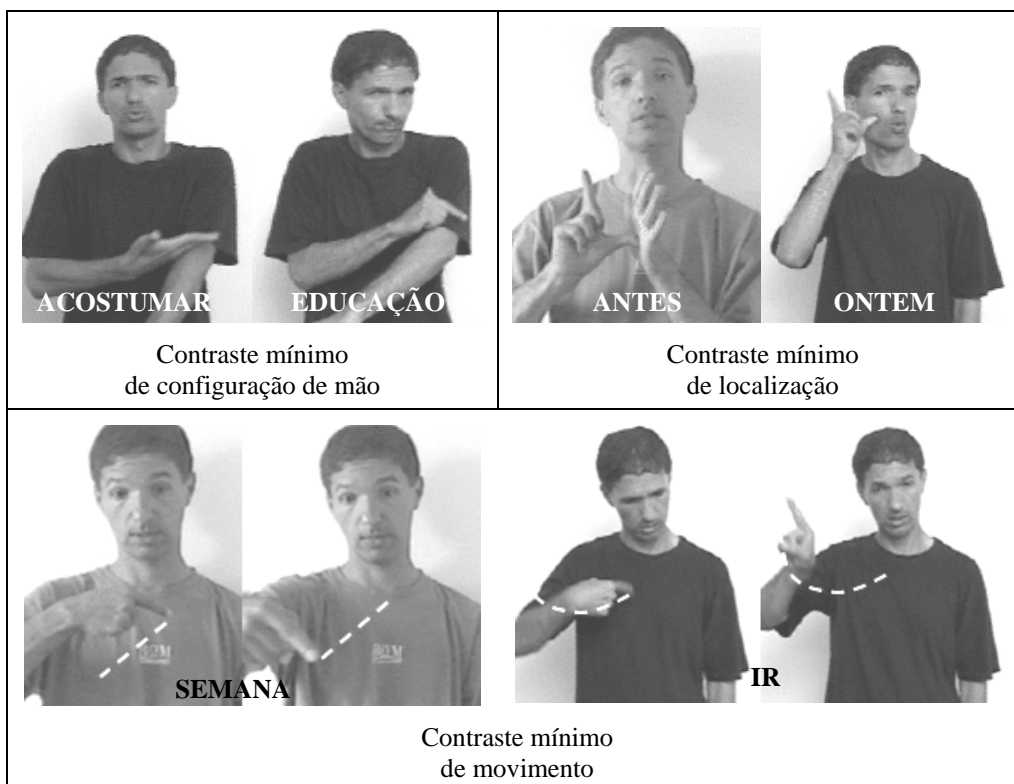


Figura 1. Exemplos de três pares mínimos na libras.

Em si, tais aspectos (i.e. uma configuração de mão na forma de “B” ou “L”; uma localização no espaço de sinalização ou ao lado da testa; um movimento retilíneo ou curvilíneo) não trazem um significado intrínseco. Contudo, combinados entre si e/ou com outros parâmetros, eles permitem a distinção de significados. Isso mostra que as LOs e as LSs se estruturam de maneira semelhante, utilizando de maneira produtiva um conjunto finito de elementos.

Apesar disso, Stokoe também apontou uma diferença entre essas duas modalidades de língua: nas LSs, os queremas pareciam ocorrer simultaneamente, ao passo que nas LOs os fonemas ocorriam em seqüencialmente. Tal diferença seria atribuída ao fato de a modalidade oral-auditiva impor uma linearidade sobre a cadeia de fala, os sons obrigatoriamente tendo que se suceder uns aos outros ao longo do tempo, diferentemente do que ocorria na composição do gesto no espaço.

O termo querema, mais tarde, acabou sendo abandonado pelos pesquisadores das LSs. Como já apontava Saussure (1970), a forma do significante refere-se a uma *imagem acústica* convencional, abstraída de realizações fonéticas concretas e infinitamente variáveis, definição que torna o conceito suficientemente abstrato para abranger não apenas representações psíquicas de sons, mas também de gestos. Apesar

desse abandono terminológico, porém, a proposta revolucionária de Stokoe foi seguida adiante, e outros pesquisadores vieram incrementar o número de aspectos dos sinais que também distinguem significado, embora não tivessem significado.

Por exemplo, Battison (1974) e Friedman (1975) identificaram a *orientação da palma da mão* – um elemento presente nas descrições de Stokoe, embora com importância secundária; Klima e Bellugi (1979) identificaram o *arranjo das mãos* (quais mãos realizam o sinal e de que modo, se ativa ou passivamente) e do *contato* (a parte da mão que toca o corpo); e Liddell e Johnson (1989) multiplicaram ainda mais os aspectos, destrinchando M em termos de *movimentos locais* (que envolvem dedos e punho) *contornos, planos e qualidade* do movimento, bem como a CM em termos de *configuração de mão, ponto de contato, face e orientação*.

Um contraste entre o que se observa nas LSs e nas LOs na produção de traços e unidades mínimas revela ainda outros paralelos e diferenças interessantes (Liddell, 2003a). Mãos e braços nas LSs podem ser entendidos como análogos ao órgão da língua nas LOs, constituindo-se nos principais articuladores móveis para produção de palavras; o espaço de sinalização nas LSs, incluindo os pontos mais produtivos do corpo, podem ser entendidos como análogos ao trato vocal nas LOs, especificando os pontos de articulação possíveis onde as palavras são realizadas; por fim, as diferentes formas de movimento nas LSs podem ser entendidas como análogas aos diferentes modos de articulação de fonemas nas LOs. Por outro lado, uma diferença marcante está no fato de que, “comparativamente falando, as LOs são um tanto limitadas nos tipos de contrastes articulatórios a elas disponíveis” considerando-se o potencial das LSs de explorar “muitas configurações distintas das mãos, ... dois articuladores ao mesmo tempo, bem como um grande número de localizações distintivas” (Liddell, 2003a: 11).

Entretanto, apesar desse diferencial, Battison (2000, publicado originalmente em 1978) demonstra que, pelo menos no que tange ao potencial de uso de dois articuladores simultaneamente, a ASL apresenta duas restrições que limitam consideravelmente as suas possibilidades articulatórias. A primeira é a *condição de dominância*, de acordo com a qual uma das mãos assume papel ativo e a outra, um papel passivo, servindo de ponto de articulação para o gesto realizado pela mão ativa (e.g. o sinal <[BANHEIRO](#)> da libras); a segunda é a *condição de simetria*, segundo a qual, no caso de as duas mãos serem ativas, ambas adquirem uma configuração de mão idêntica e movimentos

especulares (e.g. o sinal <[LÍNGUA-DE-SINAIS](#)> da libras).⁹ Apesar de algumas diferenças articulatórias entre a ASL e a libras, tais restrições foram corroboradas numa primeira análise da estrutura interna dos sinais na libras (Xavier, 2006: 107-112).

Já na década de 80, novos estudos viriam questionar a afirmação de que a estrutura interna dos sinais nas LSs seria essencialmente simultânea, em contraste com a seqüencialidade das LOs (Newkirk, 1998 [1981]; Liddell, 1984; Liddell e Johnson, 1989). Por exemplo, o sinal <[SURD@](#)> na libras envolve dois toques com o dedo indicador no rosto: um primeiro toque na região abaixo da orelha e um segundo na região ao lado da boca, intermediado por um movimento em arco no deslocamento de uma região à outra. Trocar essa ordem, tocando primeiramente a região ao lado da boca e em seguida abaixo da orelha, resulta num sinal inexistente na libras. Isso mostra que o sinal SURD@ apresenta uma seqüencialidade que não pode ser ignorada nas análises fonológicas das LSs.¹⁰

⁹ Embora Battison use a expressão “condição de *dominância*” para se referir ao caráter ativo vs. passivo dos diferentes articuladores manuais na sinalização, é importante distinguir a noção de *dominância* da noção de *atividade* (McCleary e Viotti, 2007). A mão dominante é a mão preferencialmente usada pelos surdos durante a sinalização e depende de o indivíduo ser destro ou canhoto. Na realização de sinais formados por uma mão ativa e uma mão passiva, como em BANHEIRO, por exemplo, a mão ativa tende a ser a mão dominante. Contudo, a observação inicial de dados do corpus mostra que essa preferência pode ser subvertida por razões diversas (e.g. McCleary e Viotti, 2007; ver também van der Kooij et al., 2006: 1607), de modo que, embora a correlação entre mão dominante e mão ativa seja grande, ela não deve ser assumida como absoluta.

¹⁰ Cabe destacar ainda que, no caso de alguns sinais em particular, a sua boa formação depende não apenas das mãos, mas também de sinais não-manuais envolvendo alguma parte do corpo, em especial a parte inferior do rosto. Um exemplo de palavra da libras que envolve a realização de sinais não-manuais é o sinal <[MAGR@](#)>. Tal sinal envolve o movimento da mão configurada em “I” numa trajetória retilínea de cima para baixo no espaço de sinalização, com orientação da palma voltada para fora. Porém, essa atividade manual parece exigir, obrigatoriamente, o acompanhamento de um *gesto bucal* em que as bochechas são sugadas para dentro da boca.

Gestos bucais (*mouth gestures*) são configurações da parte inferior do rosto específicas das LSs que entram na composição semântica de diversos itens lexicais e predicções. Trazendo um possível <[exemplo do corpus](#)>, há um trecho na conversa em que a sinalização de W poderia ser traduzida por algo como “Eu (ia) explicando” ou “Eu explicava”. Os sinais manuais, em si, referem-se a EU e EXPLICAR apenas, e a configuração da boca realiza a modificação aspectual durativa desse verbo.

Esses gestos bucais devem ser distintos das chamadas “imagens bucais” (*mouth pictures*) nos seguintes termos:

Tendo em vista que a comunidade surda emerge em meio a uma sociedade ouvinte, surdos passam a sua vida se defrontando com situações de línguas em contato (i.e. a sua LS em contato com a LO majoritária do país). Assim como se dá entre grupos falantes de LOs, a língua majoritária acaba se revelando mais influente sobre a língua minoritária do que o inverso, de modo que *interferências* de diversas formas do português na libras, no caso do Brasil, são comuns de serem observadas. Uma delas é a articulação de palavras da língua oral, acompanhadas ou não de som, em concomitância com a sinalização, fato que se observa inclusive em conversas entre surdos. A expressão “imagem bucal” é

A proposta de Liddell (1984) sobre a seqüencialidade dos sinais emergiu a partir de uma análise do processo morfológico de composição na ASL. Na formação de certos sinais compostos, Liddell observou que partes de um ou de ambos os sinais eram apagadas em favor da preservação de determinadas organizações seqüenciais. Um exemplo da libras que parece ilustrar esse fenômeno é o sinal <CHORAR-FATJAR>, que deve ser traduzido como “cebola”. Quando isolado, o sinal CHORAR apresenta duas seqüências bastante breves de movimento mas, ao entrar na composição, um desses movimentos é elidido.

Com base nessa análise, Liddell (1984) argumentou, então, que todo sinal fosse composto de *movimentos* (M) e/ou *suspensões* (S) – segmentos, grosso modo, comparáveis às vogais e consoantes das LOs, respectivamente. Os aspectos de Stokoe (a CM, o M e a L, assim como os demais parâmetros identificados mais tarde por outros pesquisadores), então, mudavam de estatuto: não eram mais vistos como fonemas, mas como traços distintivos que compunham segmentos fonológicos da ASL – novamente, um fenômeno comparável ao que ocorre com traços distintivos e fonemas nas LOs.¹¹

utilizada para se referir a esse fenômeno de realização de uma palavra da língua oral majoritária do país (no caso do Brasil, o português brasileiro) junto com os sinais.

Observações iniciais de um corpus mostram que as imagens bucais na fala espontânea da libras não costumam corresponder à articulação plena da palavra da língua oral, sendo muitas vezes reduzidas para alguns segmentos mais salientes dessas palavras. Esse processo de redução revela, simultaneamente, uma desvinculação da articulação bucal em relação ao português e uma subordinação dessa articulação à organização própria da língua de sinais, que por sua temporalidade e dinâmica própria impediria uma correspondência unívoca entre as palavras do português e os sinais da libras durante a sinalização fluente.

O ápice desse processo ocorre quando algumas *imagens bucais* perdem por completo o vínculo com a língua oral e tornam-se *gestos bucais*, isto é, articulações convencionais específicas das LSs. Na libras, um caso que parece ilustrar esse fenômeno é o sinal <SURD@>, cuja forma de menção parece ter incorporado (e tipificado) um segmento mais saliente da palavra oral [surdu]. Se esse processo de fato for verificado empiricamente em investigações futuras, seria possível determinar, pelo menos no que diz respeito a certos tipos de articulação da boca, um contínuo entre imagens bucais e gestos bucais – ou, em termos sociolingüísticos, um contínuo entre *interferência lingüística* e *empréstimo*.

¹¹ Sinais como CHORAR, que envolvem duas breves seqüências de movimentos idênticos, são bastante freqüentes na libras. Liddell (1984: 396) argumenta que, na ASL, os tipos de sinais existentes e os processos fonológicos e morfológicos pelos quais eles passam dão evidência de que são predominantes fonotaticamente as combinações VCVC, VC e CVC (i.e. MSMS, MS, SMS). Baseando-se no trabalho de Liddell e Johnson (1989), Xavier (2006) atesta, numa análise preliminar da libras, sinais com as seguintes possibilidades de organização segmental: S, M, SM, MS, SMS e, envolvendo repetições dessas seqüências, SMSM, MSMS, SMSSMS. Contudo, o autor não analisa a proporção de cada uma dessas possibilidades em seu corpus a fim de verificar quais combinações segmentais seriam predominantes na libras.

A proposta articulada por Liddell (1984) deu início a uma nova fase nos estudos da fonologia das LSs, tendo sido aceita em sua essência pelos pesquisadores da área a despeito de disputas teóricas envolvendo diferentes tipos de representação fonológica da seqüencialidade dos sinais (Brentari, 1998; Liddell e Johnson, 1989; Perlmutter, 1990; Sandler, 1989). Tendo em vista que o presente estudo possui um caráter descritivo, as diferenças teóricas entre essas propostas não são relevantes para a presente discussão.

2.3. Morfologia

Ao passo que a fonologia das LSs parece se constituir como um dos seus níveis de análise mais poderosos – como será discutido em maiores detalhes na seção 2.6 –, a morfologia, diferentemente, parece ser um nível de análise significativamente limitado quando comparado com línguas como o português.¹² A grande maioria dos sinais na ASL, e possivelmente também na libras, são monomorfêmicos. Sinais complexos existem, mas estão reduzidos a basicamente três tipos: a) sinais compostos (e.g. Liddell, 1984); b) sinais com incorporação de números (e.g. Frishberg e Gough, 2000 [1973]); e c) sinais modificados aspectualmente (e.g. Klima e Bellugi, 1979). Nenhuma dessas formas é marcada por meio de afixação, como é comum de ocorrer nas LOs ricas morfologicamente.¹³

Sinais complexos formados por composição já foram exemplificados na discussão sobre fonologia acima. Liddell (1984) especifica duas regras de formação que, na ASL, resultam na alteração dos sinais integrantes da composição em relação a cada um dos sinais individuais: a *regra da suspensão em contato* diz que, se um sinal tem um segmento do tipo suspensão em que a mão entra em contato com o corpo, quando integrar uma composição esse sinal envolverá apenas esse segmento de suspensão, elidindo os demais; já a *regra da seqüência única* diz que repetições de seqüências de movimento e suspensão serão eliminadas quando o sinal integra uma composição. O sinal composto <[CASA-ESTUDAR](#)>, que poderia ser traduzido como

¹² Evani Viotti, comunicação pessoal.

¹³ Na ASL, Liddell (2003a) reporta dois casos isolados de sinais que são formados por um sufixo, e alguns poucos casos de prefixos relacionados à expressão de valores numéricos. De modo geral, porém, o processo de afixação não parece ser produtivo nas LSs até hoje estudadas.

“escola”, é um exemplo que ilustra a operação de ambas as regras, o que sugere que elas estejam presentes também na libras.¹⁴

O segundo tipo importante de sinais morfológicamente complexos são aqueles que apresentam incorporação de numerais. Na libras, os numerais de 1 a 9 são expressos por sinais monomorfêmicos e, ao que parece, o UM, o DOIS, o TRÊS e o QUATRO são incorporados produtivamente em outros itens lexicais, tais como aqueles referentes a duração em horas (e.g. <[HORA-DOIS](#) e [HORA-TRÊS](#)>), número de pessoas (e.g. <[PESSOA-UM-ANDAR](#) e [PESSOA-QUATRO-ANDAR](#)>), valores monetários (e.g. <[REAL-DOIS](#) e [REAL-TRÊS](#)>),¹⁵ e principalmente calendário (e.g. <[MÊS-UM](#), [SEMANA-DOIS](#) e [SEMESTRE-TRÊS](#)>). Além disso, vários pronomes apresentam também esse tipo de incorporação (e.g. <[EL@-DOIS](#) e [NÓS-TRÊS](#)>). Em todos esses casos, diferentemente das LOs, a incorporação morfêmica ocorre de maneira simultânea, e não por meio de uma afixação sequencial (Liddell, 2003a: 19).

O terceiro e último tipo de sinal complexo na libras são os sinais predicativos modificados aspectualmente. Um processo bastante produtivo de modificação de aspecto nas LSs é a *reduplicação*, também documentada principalmente nas línguas indígenas e asiáticas. Por exemplo, na libras, a reduplicação das seqüências de movimento e suspensão acompanhada de sinais não-manuais pode expressar a idéia de iteratividade (e.g. <[ESTUDAR](#) e [ESTUDAR-MUIT@](#)>). Além disso, a expressão de diferentes aspectos parece ser diferenciada formalmente por mudanças na qualidade dos movimentos reduplicados, e pelo possível acompanhamento de sinais não-manuais específicos (Klima e Bellugi, 1979; Liddell, 2003a).

Os estudos das LSs nesta área mostram que os vários tipos de modulações observados nos sinais ocorrem tanto por motivações fonológicas, que não acarretam

¹⁴ Duas observações, uma empírica e uma teórica, devem ser feitas sobre esse exemplo. Em primeiro lugar, o sinal utilizado para se referir à “escola” na libras muitas vezes envolve apenas o sinal ESTUDAR. Sem uma pesquisa ampla, tanto sincrônica quanto diacrônica, não é possível saber em que medida o uso do sinal ESTUDAR para se referir à “escola” originou-se por meio de uma redução do composto CASA-ESTUDAR no processo de evolução da língua, ou como uma extensão metonímica de “estudo” para “o local onde se estuda”. Em segundo lugar, os exemplos que Liddell (2003) oferece sobre a operação dessas duas regras não envolvem composições em que ambos os sinais integrantes são formados por repetições de seqüências de movimentos e suspensões, como é o caso de CASA e ESTUDAR. No composto da libras, nota-se que a eliminação das seqüências repetidas se dá com apenas um dos sinais integrantes (no caso, CASA), e não com ambos, como a regra parece sugerir.

¹⁵ No caso de valores monetários, a incorporação parece não estar restrita aos numerais UM a QUATRO, podendo incluir outros números também.

mudanças de sentido e estão possivelmente ligadas à prosódia (e.g. fusão de palavras compostas por meio da elisão de segmentos), quanto por motivações morfológicas, que envolvem mudanças de significado (e.g. modificações de segmentos para expressão de nuances aspectuais). As seções de análise desta pesquisa irão tratar das modulações do primeiro tipo, isto é, de natureza fonológica/prosódica.

2.4. Sintaxe

Os estudos da sintaxe da ASL e das demais LSs do mundo ganharam um novo impulso a partir da década de 70. Até então, a maior parte dos pesquisadores acreditava que a ordem das sentenças na ASL era basicamente livre, sem restrições, tendo em vista que, em diferentes contextos discursivos, os sinais correspondentes a sujeito e objeto apareciam posicionados de diferentes maneiras em relação ao verbo. Essa visão começou a mudar a partir de estudos que vieram destacar o importante papel dos sinais não-manuais, principalmente relativos ao rosto e à cabeça, na identificação de fenômenos sintáticos (Baker, 1976; Baker e Padden, 1978; Liddell, 1978).

Em Liddell (1978), uma história repleta de personagens sem nome foi elaborada a fim de se verificar como orações relativas eram produzidas na ASL. Foi solicitado a alguns surdos que lessem e memorizassem a história para, em seguida, recontá-la em ASL. Ao notar que as expressões faciais dos narradores pareciam estar desempenhando um papel importante, Liddell tirou fotos da tela do televisor em cada sinal isolado. Com isso, ele pôde constatar que os sinalizadores mantinham uma expressão facial e uma posição de cabeça particular durante todo o período em que realizavam os sinais relacionados à relativa, configuração essa que mudava tão logo uma nova predicação era iniciada, como mostra a figura 2 abaixo – reproduzida de Liddell (2003a: 54).



Figura 2. Expressão facial marcando e delimitando relativa na ASL.

No início dos estudos das ASL, Stokoe et al. (1965) já haviam destacado o fato de que algumas expressões faciais desempenhavam um papel importante, argumentando que, para que questões do tipo sim/não fossem reconhecidas como tais, elas necessariamente precisavam ser acompanhadas de uma expressão facial e posição de cabeça marcadas. Mais tarde, Bellugi e Fisher (1972) e Baker (1976) apontaram ainda o papel do balanço da cabeça de um lado para o outro acompanhado de uma expressão facial não-neutra na realização de orações negativas.

Wilbur (2000) propõe que as regiões superior e inferior do rosto, relativas aos sinais não-manuais, estejam relacionadas a diferentes domínios sintáticos: sinais da parte superior do rosto ou a cabeça (sobrancelha, olhar, posição e movimentos de cabeça) ocorreriam com constituintes maiores, tais como orações e sentenças; sinais da parte inferior do rosto (boca, língua, bochechas), diferentemente, se associariam com itens lexicais ou com os sintagmas em que tais itens aparecem, em especial para a veiculação de informações adjetivais ou adverbiais (p. 224-225).

Com o estudo de Liddell (1978) sobre as relativas, então, passa a ganhar força a proposta de que sinais não-manuais, formados principalmente por determinadas combinações de posição da cabeça e expressão facial, serviriam para delimitar constituintes gramaticais como sintagmas e orações de maneira geral. Baker e Padden (1978), então, corroborando a relevância dessa correlação por meio de experimentos com condicionais, argumentam explicitamente a favor da necessidade de análise dos sinais não-manuais em investigações voltadas à delimitação dos constituintes básicos da ASL.

Tais estudos levaram muitos pesquisadores a rever a questão da ordem dos constituintes na sentença, até então assumida como livre na ASL. Em especial, a descoberta de dois sinais não-manuais – um específico para a marcação de tópicos, formado por uma combinação de posição da cabeça e expressão facial, e outro para marcação de *tag questions*, formados por um pronome acompanhado de um aceno de cabeça – levou Liddell a afirmar que a ordem da oração principal não seria variável. Ao contrário, a distribuição de sujeito e objeto seria bastante estrita e a aparente variabilidade decorreria de dois fatores: os sujeitos e objetos poderiam ser omitidos da oração principal quando já se mostrassem proeminentes no discurso precedente; e a oração como um todo poderia ser antecedida por tópicos e/ou sucedida por *tags*

pronominais quando houvesse necessidade de salientar certos referentes do discurso (Liddell, 1980).¹⁶

2.5. Prosódia

Estudos sobre a prosódia das LSs, diferentemente dos demais níveis de análise, demoraram mais tempo para se estabelecer no campo. Uma das possíveis razões está no fato de que, como esses estudos vêm demonstrando, a prosódia nessas línguas é em grande medida veiculada por meio de sinais não-manuais, que começaram a receber uma atenção mais cuidadosa dos lingüistas por volta da década de 80.

Dentre as funções atribuídas à prosódia, a que mais diretamente interessa ao presente estudo é a da segmentação gramatical do discurso por meio da imposição de agrupamentos prosódicos.¹⁷ Uma das formas mais simples e diretas de identificação desses agrupamentos é a pausa, que nas LOs corresponde à ausência de vocalização. Nas LSs, uma maneira mais óbvia de se identificar pausas é observando o momento em que as mãos retornam do espaço de sinalização para uma posição de repouso. Winston (2000) argumenta, porém, que além dessa situação mais óbvia, manifestações de pausa nas LSs podem ser “mais complexas”, envolvendo períodos em que o falante suspende um sinal no ar, sem movimento, por um período relativamente prolongado de tempo. Nespor e Sandler (1999) acrescentam que essa manutenção da suspensão, em geral, envolve um relaxamento da configuração e localização das mãos (p. 17). O fator crucial, nesse sentido, seria a *ausência de movimento*, em contraposição à ausência de som nas LOs.¹⁸

¹⁶ No Brasil, a sintaxe da libras tem sido estudada principalmente sob a perspectiva da gramática gerativa, em trabalhos como Quadros (1999) e Quadros e Karnopp (2004).

¹⁷ Ainda que diversos estudos das LSs tragam contribuições para essa questão, a maioria o faz de uma maneira indireta (e.g. análise de relativas, em Liddell, 1978; de delimitação de tópicos no discurso, em Winston, 2000; de seqüências rítmicas, em Boyes-Braem, 1999; de contrastes semânticos, em Wibur e Patschke, 1998). Apesar disso, a presente seção irá discutir esses trabalhos em termos de sua relevância para a questão da identificação de agrupamentos prosódicos, que interessa mais diretamente a esta tese.

¹⁸ Winston propõe uma diferença entre três tipos de pausa na ASL, que ela chama de “preenchida”, “prosódica” e “extra-lingüística”. A *pausa preenchida* seria identificada pela manutenção de uma suspensão do sinal no espaço; a *pausa prosódica*, por mudanças em vários parâmetros corporais (mudanças de configuração do tronco, cabeça, etc) de uma unidade para outra; e a *pausa extra-lingüística*, por períodos de inatividade das mãos, quando o falante demonstra estar planejando a sua fala.

Tais categorias apresentam, a meu ver, o problema de não serem claramente distinguíveis umas das outras (e.g. o planejamento da fala – “extra-lingüístico” – pode envolver também sinais mantidos em

Um segundo critério importante na delimitação de agrupamentos prosódicos nas LOs é o alongamento final, frequentemente correlacionado ao limite de unidades gramaticais. Liddell (1978), em sua análise das orações relativas na ASL, assinala que os sinais que aparecem em posição final nessas orações apresentam uma duração significativamente maior do que os mesmos sinais em posição inicial ou medial na oração. Resultados similares foram encontrados em estudos subseqüentes da ASL, agora especificamente voltados à análise prosódica (e.g. Coulter, 1993; Wilbur e Nolen, 1986).

Um terceiro critério importante na delimitação de agrupamentos prosódicos é a definição de um contorno entoacional coeso, que nas LSs têm sido relacionados às expressões faciais. Por exemplo, enquanto perguntas sim/não e qu- nas LOs são acompanhadas de uma prosódia marcada, nas LSs elas são marcadas por expressões faciais específicas. Nespor e Sandler (1999: 27) argumentam, nesse sentido, que uma diferença de modalidade na expressão da entoação possa ser apontada: nas LOs as melodias tonais são produzidas por um único articulador (i.e. as cordas vocais), recaindo sobre sílabas e se dispondo seqüencialmente na cadeia da fala; nas LSs, diferentemente, as “melodias” faciais seriam produzidas pela configuração simultânea de vários articuladores (e.g. sobrancelha, pálpebra, boca, cabeça) que recairia igualmente sobre todos os sinais de um agrupamento prosódico.¹⁹

suspensão no ar – “preenchida”; a ausência de movimento – “preenchida e extra-linguística” – pode estar relacionada não apenas com o planejamento mas também com a delimitação gramatical – “prosódica”).

¹⁹ Dentro do modelo teórico da fonologia prosódica em que as autoras trabalham (Nespor e Vogel, 1986), o domínio prosódico se organiza, assim como a sintaxe, de maneira hierárquica. Partindo do nível mais baixo para o mais alto na hierarquia, estão as seguintes categorias: a sílaba, o pé, a palavra fonológica, o grupo clítico, a frase fonológica, a frase entoacional e o enunciado fonológico. Para Nespor e Sandler (1999), cujo estudo toma por base a língua de sinais israelense (ISL), as expressões faciais têm escopo, especificamente, sobre a *frase entoacional*.

Sandler (1999) aponta ainda o papel de dois outros recursos da ISL na delimitação de determinados níveis da constituência prosódica. Um desses recursos é [<o espraçamento da mão não-dominante dentro dos limites da frase fonológica>](#). (Sandler, 1999: 24-25). Um segundo recurso, similar ao que ocorre com as LOs, é a [<cliticização de uma palavra funcional fraca>](#) (pronomes dêiticos articulados com uma mão) em uma palavra de conteúdo (articulada com duas mãos), resultando numa ressilabificação, de dois movimentos independentes em um único, o que permitiria a identificação de *palavras fonológicas* (Sandler, 1999: 19). Esse processo de cliticização é corroborado pela observação de que, na realização da palavra fonológica formada por dois sinais distintos, [<o sinalizador produz uma única imagem bucal>](#), correspondente apenas ao sinal mais forte, de conteúdo, sem referência ao sinal fraco, funcional. Segundo a autora, esse fenômeno de cliticização, embora viole a condição de simetria como princípio de boa formação dos sinais (cf. seção 2.2), constitui-se num processo fonológico amplamente atestado nas línguas do mundo.

Ainda contribuindo para a identificação de agrupamentos prosódicos nas LSs aparecem na literatura o olhar, a pálpebra, e a exploração do espaço. Engberg-Pederson (1993) e Boyes-Braem (1999), por exemplo, apontam que constituintes gramaticais podem ser distinguidos uns dos outros por mudanças na alocação espacial dos sinais correspondentes a cada constituinte (e.g. o falante sinaliza uma frase à esquerda e outra, à direita). Focalizando no uso de sinais não-manuais, Baker e Padden (1978) apontam indícios de que tanto a mudança de direção do olhar quanto as piscadas dos sinalizadores em conversas entre surdos estejam correlacionadas às fronteiras de unidades gramaticais.

Wilbur (1994), concentrando-se especificamente na análise das piscadas, argumenta que, na literatura sobre o sistema ocular, tal fenômeno se mostra relacionado a três funções básicas: piscadas de reflexo, piscadas involuntárias periódicas (para lubrificação dos olhos) e piscadas voluntárias. Ao confirmar a proposta de Baker e Padden (1978) de que piscadas recorrem ao término de possíveis constituintes gramaticais na ASL, a autora propõe, então, que tais piscadas sejam movimentos fisiológicos involuntários que acabam submetidos à organização da língua. Nespor e Sandler (1986) reforçam essa proposta, fazendo uma analogia entre a necessidade de piscadas durante a produção de fala sinalizada, e a necessidade de inspiração de ar durante a produção de fala oral, responsável pela ocorrência de pausas nas fronteiras entre unidades gramaticais.

Além das expressões faciais, do olhar, das piscadas e das modulações de sinais manuais, outro recurso das LSs que pode ser relacionado à delimitação de agrupamentos prosódicos é a inclinação do corpo, que envolve o tronco como um todo ou os ombros apenas. De acordo com Wilbur e Patschke (1998), movimentos do tronco para frente e para trás aparecem associados à expressão de vários tipos de contrastes: envolvimento *vs.* não-envolvimento, inclusão *vs.* exclusão e afirmação *vs.* negação. Em alguns casos, esses sinais não-manuais aparecem como parte convencional de sinais cujo conteúdo semântico envolve uma dessas oposições (e.g. ENVOLVER *vs.* HESITAR na ASL); em outros, eles aparecem associados a frases semânticamente opostas.

Concentrando-se sobre um tipo específico de foco contrastivo dentre os vários analisados por Wilbur e Patschke (1998), van der Kooij et al. (2006) destacam a possibilidade de, na língua de sinais holandesa, haver uma preferência pela inclinação do tronco para a direita e esquerda, ao invés de para trás e para frente, acompanhada da

alternância da mão direita ou esquerda como mão ativa na sinalização.²⁰ Ainda sobre esses movimentos laterais de tronco, Boyes-Braem (1999) aponta evidências, na língua de sinais suíço-germânica, de que tais movimentos revelem padrões temporais rítmicos ao longo do discurso, e que a oscilação para um lado e para o outro esteja correlacionada a trechos maiores da narrativa.

Tomados em conjunto, esses estudos revelam que parte substancial da prosódia nas LSs se manifesta por meio de sinais não-manuais, ainda que modulações tais como o alongamento final também se mostrem relevantes. Embora haja uma tendência, na literatura em questão, de se associar os sinais não-manuais (rosto, cabeça e tronco) com o nível supra-segmental, e os sinais manuais com o nível segmental, parece mais prudente assumir que a prosódia nas LSs deva envolver também as mãos/braços. Na verdade, a aparente idiosincrasia das LSs no uso que fazem dos sinais não-manuais (i.e. rosto, cabeça e tronco) poderia ser pelo menos relativizada quando nos damos conta da desconsideração histórica desses mesmos recursos na análise da prosódia das LOs (mas ver os vários trabalhos de Dwight Bolinger; Loehr, 2004; McNeill, 1992; e, em campos de investigação mais recentes, os estudos sobre a *prosódia visual*, e.g. Barkhuysen et al., 2008; Krahmer e Swerts, 2007; Swerts e Krahmer, 2006; e sobre o reconhecimento da fala em estudos multimodais, e.g. Quek et al., 2001, 2002).

2.6. Gestualidade

Até os estudos de Stokoe (1960), a própria lingüística não havia escapado do senso comum no modo como enxergava o meio de comunicação dos surdos. O que hoje chamamos “língua de sinais” era antes tido como uma forma de linguagem universal, icônica e/ou pantomímica, sem o tipo de estruturação que sabemos ser característico das línguas humanas. Desde os estudos de Stokoe, então, um esforço considerável por parte dos pesquisadores das LSs tem sido feito no sentido de demonstrar que essas línguas, assim como as LOs, compartilham as propriedades básicas das línguas naturais, como a produtividade e a arbitrariedade (Saussure, 1970).

²⁰ Ambos os trabalhos acima citados destacam que os movimentos de tronco devem ser analisados não em termos absolutos, mas relativos à posição do corpo na unidade prévia e na unidade subsequente, sejam elas sinalizadas pelo próprio falante, como argumentam Wilbur e Patschke (1998), ou mesmo pelo seu interlocutor, como argumentam van der Kooij et al. (2006).

Pode-se dizer que esse esforço não foi em vão e que, hoje, o estatuto lingüístico das LSs já se mostra amplamente aceito, pelo menos dentro da comunidade lingüística. Como visto nas seções acima, que trazem uma amostra pequena porém ilustrativa do que tem sido feito na área, os pesquisadores das LSs foram capazes de demonstrar de que maneira os diferentes níveis de análise que integram o estudo das LOs podem se manifestar em línguas de modalidade distinta. Tal demonstração, contudo, não esteve livre de excessos. No esforço de conferir estatuto científico às LSs, algumas características patentes do uso dos sinais, tais como a gradiência, a iconicidade e a motivação foram varridos para debaixo do tapete, em favor de análises que valorizavam a discrição e a arbitrariedade típicas das gramáticas normativas e descritivas tradicionais das LOs.

É interessante notar, nesse sentido, que o questionamento sobre o caráter puramente discreto e arbitrário da gramática das LSs não tenha partido do próprio campo, tendo sido impulsionado pelo desenvolvimento de teorias emergentes no âmbito das próprias LOs, em análises sobre a relação entre língua e gesto (e.g. Kendon, 1980; McNeill, 1992) e língua e cognição (e.g. Langacker, 1987, 1991; Lakoff e Johnson, 1980; Lakoff, 1987; Fauconnier, 1985).²¹ Dessas teorias emergem algumas lições que tiveram um impacto profundo sobre os estudos das LSs: a) a gestualidade é parte integrante do uso vivo da língua e revela-se intimamente relacionada aos aspectos prosódicos e semânticos da fala; b) a arbitrariedade do signo não implica uma ausência de motivação, mas sim o papel da convenção sempre seletiva que cada comunidade lingüística faz de sua experiência; e c) todo o nosso conhecimento abstrato (incluindo o gramatical) é construído sobre um conhecimento mais primitivo e concreto que, por sua vez, é construído a partir de nossa interação corporal e social com o mundo.

Livres dos mitos e preconceitos sobre a relação entre língua e gesto, e livres da responsabilidade de ter que demonstrar que as LSs eram iguais às LOs em todos os seus níveis de análise, muitos pesquisadores da área iniciaram então uma investigação séria sobre as possíveis diferenças entre LSs e LOs, e os resultados têm trazido uma nova luz sobre problemas que permaneceram por muito tempo obscuros na área – além de contribuir com uma nova perspectiva sobre questões relacionadas às próprias LOs.

²¹ Mas ver os trabalhos de Stokoe, que, ao longo de sua carreira, parece ter sempre buscado dar conta da dimensão gestual e icônica dos sinais sem com isso abdicar dos princípios lingüísticos (e.g. Armstrong, Stokoe e Wilcox, 1995; Stokoe, 2001).

Como consequência, hoje dispomos de uma perspectiva renovadora sobre a fonologia, a morfologia, a sintaxe e prosódia das LSs, que pode ser contraposta a um olhar sobre a gramática marcadamente enviesado pela cultura ocidental, culta e letrada.

Na fonologia, a análise revolucionária de Stokoe sobre a capacidade recombinativa de unidades mínimas dos sinais permaneceu em sua essência inalterada, com exceção das reformulações necessárias para dar conta da seqüencialidade dos sinais e, conseqüentemente, dos processos fonológicos e morfológicos observáveis. Contudo, alguns pesquisadores têm hoje questionado a natureza dessas unidades mínimas, ou traços, que entram na composição do sinal: seriam eles, de fato, *destituídos de sentido*, embora capazes de distinguir sentidos?

A figura 3 abaixo mostra três sinais da libras: <ÁRVORE>, <PENSAR> e <FATIAR>.²² Olhando-os, é impossível negar a sua dimensão icônica. Em ÁRVORE, podemos identificar o solo (representado pela mão passiva), o tronco (representado pelo antebraço da mão/braço ativo) e a copa (representada pela mão ativa), o que revela a alta iconicidade da *configuração das mãos* e da sua *disposição espacial*. Em PENSAR, podemos identificar, pelo lugar em que o sinal é realizado, a região à qual atribuímos o ato de pensar (representada pela própria cabeça), o que revela a alta iconicidade do *ponto de articulação*. Em FATIAR, podemos identificar o objeto cortante (representado pela configuração da mão ativa em “B”), o objeto cortado (representado pela mão passiva em “C”)²³ e a ação de cortar (representada pelo movimento da mão ativa, em relação à mão passiva), o que revela a alta iconicidade não apenas da configuração de mão, mas também do *movimento* do sinal e do *ponto de articulação* (a lateral da mão passiva).

²² Exemplos da libras nesta seção foram tirados de dados do corpus do grupo ECS e do *dicionário libras on line*, disponível na rede no seguinte endereço: <http://www.dicionariolibras.com.br/website/dicionariolibras/dicionario.asp?cod=124&idi=1&moe=6&categoria=66>.

²³ Configuração de mão em “B”:  e configuração de mão em “C”: .



Figura 3. Sinais da libras com alto grau de iconicidade.

Esse processo de construção de sinais fazendo uso de recursos icônicos é altamente produtivo nas LSs, e está também presente nas LOs, embora de maneira bem mais limitada (Taub, 2000).²⁴ Taub propõe, então, um “modelo de construção analógico” para dar conta desse fenômeno produtivo, que envolve três dimensões distintas: seleção de imagem, esquematização e codificação. O esquema apresentado na figura 4 abaixo, reproduzido de Taub (2000: 35), ilustra a aplicação do modelo proposto pela autora à formação do sinal **ÁRVORE** da ASL.

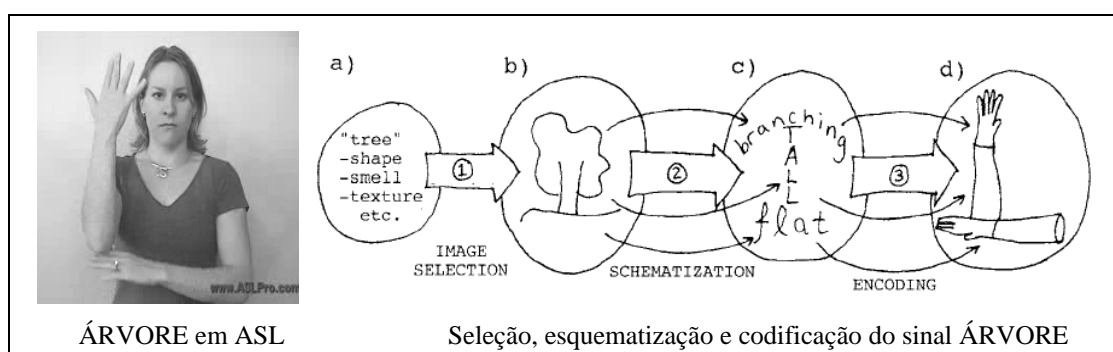


Figura 4. Ilustração do modelo de construção analógico de Taub.

Segundo Taub, nosso conceito de “árvore” carrega todo o nosso histórico (subjetivo) de interação com árvores, podendo envolver o conhecimento de suas formas, cheiros, texturas, sons que produzem, além de todo conhecimento enciclopédico que

²⁴ Nas LOs, a iconicidade tem sido investigada sob a rubrica de conceitos tais como os de onomatopéia, simbolismo sonoro, ideofone, entre outros, constituindo-se num fenômeno de destaque, em particular, em várias análises gramaticais no âmbito da lingüística cognitiva. Para Taub (2000), a maior iconicidade das LSs em comparação com as LOs pode ser explicada da seguinte maneira: “Primeiramente, nós possuímos mais imagens motoras e visuais associadas a conceitos do que sonoras – por exemplo, não há qualquer som característico para a categoria *mesa*, e no entanto há um formato característico. Em segundo lugar, a modalidade sinalizada, com o seu uso de movimentos corporais, expressões faciais, configurações das mãos e dos braços, e o espaço próximo ao sinalizador, possui um grande número de alternativas de construir análogos lingüísticos para as imagens mentais. A modalidade oral tem pouco mais do que a ordenação dos sons e o *pitch* da voz do falante” (p. 38).

pode ser adquirido indiretamente sobre elas. De todo esse potencial, é feita uma *seleção* metonímica de um ou mais traços imagéticos para serem representativos do conceito na língua. No caso da ASL – e, como veremos mais adiante, também da libras –, a imagem selecionada é a de uma árvore que nasce perpendicularmente ao chão, com um tronco liso e uma copa, sem referência, por exemplo, à raiz, folhas, frutos, etc. Essa escolha é arbitrária, embora motivada. Outras LSs, por exemplo, escolhem outros traços para essa representação, como a língua de sinais húngara, que seleciona apenas o solo e o tronco; a língua de sinais turca, que seleciona apenas a copa; e a língua de sinais chinesa, que seleciona apenas o tronco – como mostra a figura 5 abaixo.²⁵



Figura 5. Arbitrariedade e motivação na seleção de imagens do sinal ÁRVORE nas línguas de sinais húngara, turca e chinesa.

No processo de *esquematisação*, detalhes vívidos inerentes às instâncias são abstraídos em favor de elementos abstratos e simplificados (i.e. esquemáticos). Na figura 4 acima, por exemplo, a esquematização de ÁRVORE na ASL mostra que detalhes sobre o solo (e.g. se é liso ou esburacado), o relevo (e.g. se é montanhoso ou rochoso), o tronco (e.g. se possui várias ramificações, se é grosso ou fino) e a copa (se é ampla ou mirrada, alta ou baixa) cedem lugar a apenas dois elementos: superfície plana e objeto comprido vertical com a ponta alargada. Apesar dessa esquematização, são mantidos elementos estruturais suficientes para que os mapeamentos entre o domínio mais abstrato (o sinal da ASL) e o domínio mais concreto (a imagem da árvore) sejam realizados.

²⁵ Os exemplos foram tirados de <http://www.hallatlan.hu/en/> (língua de sinais húngara), <http://turkisaretdili.ku.edu.tr/en/wordlist.aspx#> (língua de sinais turca), e Klima e Bellugi (1979) (língua de sinais chinesa). Klima e Bellugi foram os primeiros a apontar a natureza arbitrária, embora motivada, dos sinais nas LSs.

Uma última dimensão do processo de formação do sinal descrito por Taub, então, é a *codificação*. Já na esquematização, recortes importantes são feitos diante da riqueza de detalhes dos referentes, a fim de que o conceito possa ser representado sem uma violação dos princípios de economia que toda língua exige. Na codificação, esse recorte assume um caráter ainda mais particular, uma vez que cada língua irá trazer um repertório convencional distinto de formas potencialmente utilizáveis para a representação lingüística do conceito.

Comparando as convenções adotadas nas línguas de sinais italiana, croata e brasileira, por exemplo, é possível observar uma série de distinções. Na língua de sinais italiana, como se vê na figura 6 abaixo, o sinal é bastante similar ao da libras, exceto pela configuração da mão que representa a copa e pelo fato do sinal não envolver qualquer movimento. O sinal da libras e da língua de sinais croata possuem as configurações de mão iguais e ambos possuem movimento; porém, há uma inversão em termos de qual mão terá papel ativo ou passivo e uma diferença no tipo de movimento realizado: na libras, a mão/braço que representa o tronco e a copa fazem um movimento local rotatório com giro de 180°; na língua de sinais croata, a mão que representa o tronco e a copa permanece passiva enquanto a outra mão, ativa, se desloca no espaço num movimento retilíneo até atingir a “base do tronco”.^{26,27} Além disso, a mão ativa na língua de sinais croata apresenta uma orientação da palma inversa à observada na língua de sinais italiana e na brasileira.

²⁶ Essa diferença de movimento pode sugerir uma diferença entre a língua de sinais croata e a brasileira também no processo de seleção de imagem. Na língua de sinais croata, o movimento e a orientação da palma da mão ativa sugerem, iconicamente, que a imagem selecionada para representar a árvore seja a de um objeto cortante, como por exemplo um machado, golpeando a base da árvore. A abertura na interpretação quanto à fonte de iconicidade dos sinais poderia ser investigada, por exemplo, com um estudo etimológico da palavra e/ou um estudo da relação histórica dessa língua com outras LSs.

²⁷ Os exemplos foram tirados de <http://www.truveo.com/Albero/id/2619326486> (língua de sinais italiana), <http://www.crodeafweb.net/rjecnik/english/indexeng.html> (língua de sinais croata).



Figura 6. Arbitrariade e motivação na codificação de elementos do sinal ÁRVORE na língua de sinais italiana, croata e brasileira.

Como destaca Taub (2000), reconhecer essa iconicidade não implica equivaler as LSs à mímica ou pantomima, uma vez que “os itens lingüísticos são forçados a adequar-se às categorias fonéticas e semânticas da língua, enquanto a mímica é restrita apenas pelo poder de conceitualização do imitador e suas habilidades físicas” (p. 37).²⁸ Com essa análise sobre a iconicidade dos traços mínimos que compõem a fonologia das LSs, então, a autora dá conta da arbitrariedade do signo, um princípio fundamental da lingüística, sem negligenciar a sua dimensão icônica.²⁹

²⁸ Essa iconicidade dos traços mínimos que compõem os sinais das LSs tem conduzido alguns autores a achados que, embora reveladores, me parecem interpretados de maneira inadequada. Por exemplo, Capovilla e Raphael (2005) descobriram um interessante repertório de padrões icônicos produtivos na formação dos sinais da libras (e.g. a abertura de mãos e dedos para representar a propagação de um “objeto” físico ou abstrato (como em ESPALHAR, ACENDER, PROJETAR, ENSINAR, entre outros), mas interpretam esses padrões como exemplo do que eles denominam “morfemas metafóricos”.

²⁹ A discussão de Taub prossegue com uma análise de sinais com significado abstrato, tal como ANALYZE na ASL. A autora mostra que, nesses casos, deve ser considerado, previamente à “seleção de imagem”, um processo inicial de metaforização do conceito abstrato em termos de um conceito mais concreto, nas linhas da análise de Lakoff e Johnson (1980).

Por exemplo, uma metáfora comum em nossa cultura ocidental descrita por Reddy (1979) é o que ele chamou de “metáfora do conduíte” segundo a qual, *idéias são coisas que existem em nossas cabeças/mentes e o processo de comunicação envolve a transmissão dessas coisas de uma mente para a outra*. Esse mapeamento de um domínio mais abstrato, de comunicação, em termos de um domínio mais concreto, de troca de objetos de lugar, permeia o uso mais cotidiano que fazemos da língua (e.g. “quem se colocou essa idéia na cabeça?”, “eu dei uma idéia pra ele de como queria a casa”, “a gente gosta de trocar idéias”, etc).

Aplicando a proposta de Taub aos sinais <COMUNICAR e NÃO-SE-ENTENDER> da libras, num exercício de reflexão, podemos sugerir que, em ambos os sinais, as mãos representam as idéias (i.e. objetos) e o movimento representa o processo de troca dessas idéias. No sinal COMUNICAR, em que se assume a princípio o caráter bem sucedido da “troca”, as mãos se deslocam de maneira alternada entre o corpo do sinalizador e o espaço a frente, num movimento que representa iconicamente essa troca; já no sinal NÃO-SE-ENTENDER, as idéias representadas pelas mãos encontram uma barreira em meio à transferência de lugar, chocando-se uma com a outra antes de alcançar o seu destino final. Para ver uma discussão mais detalhada sobre esse processo no caso de sinais abstratos da ASL, ver Taub (2000: 39-43).

Embora possa parecer óbvio, vale destacar que a análise de Taub (2000) sobre a iconicidade dos sinais

O incrível potencial de representação icônica nas LSs parece poupar essas línguas da necessidade de construir sentidos por meio de recursos e processos morfológicos e sintáticos. Esse talvez seja um dos motivos pelos quais esses dois domínios sejam relativamente simples nas LSs, quando comparados às LOs de maneira geral. Essa rica exploração icônica da gestualidade e do espaço imediato para a veiculação de sentidos nas LSs tem servido de base de explicação para dois outros importantes fenômenos relativos a esses âmbitos de análise: os verbos indicadores e os verbos descritivos (Liddell, 2003a).

Verbos indicadores são também conhecidos como “verbos direcionais”, “verbos de concordância”, “verbos de flexão”, entre outros. Trata-se de verbos que indicam o sujeito e o objeto da ação pela maneira como o seu movimento e orientação são realizados no espaço. Por exemplo, o sinal <CONTAR>, em sua forma de menção, parte do corpo do sinalizador em direção ao espaço neutro. Contudo, essa forma nunca é encontrada na sinalização natural, em que “aquele que conta” e “aquele para quem algo é contado” devem obrigatoriamente ser indicadas por meio do direcionamento do sinal no espaço. No caso de os referentes do sujeito e objeto estarem presentes no ato de enunciação, o ponto de partida e de chegada do sinal levará essa presença física dos referentes do verbo em consideração. No caso de o referente estar ausente, um local no espaço será utilizado para localizá-lo, e, a partir daí, esse local será utilizado como ponto de referência para o direcionamento desses verbos.

Sob a perspectiva linguística majoritária no campo das LSs, de acordo com a qual toda fonte de significado provém de morfemas, verbos dessa natureza são considerados verbos de concordância. A proposta é a de que o verbo concorde com o sujeito e o objeto por meio da alteração do ponto de articulação inicial e final do sinal, cada um deles considerado um “morfema de concordância”. Problemas nesse tipo de

não deve ser tomada como um indicador de que todos os sinais das LSs possam ser explicados nesses termos. Em primeiro lugar, essa análise não nega o fato de que possam haver sinais totalmente arbitrários. Em segundo lugar, já foram documentados exemplos em que a iconicidade original dos sinais é perdida no processo de evolução da língua (Klima e Bellugi, 1979). Uma possibilidade é a de que a imagem de um referente que um dado sinal representa deixe de ser auto-evidente com o passar dos anos pelas modificações que esse referente sofre no mundo objetivo, como ocorre cotidianamente com as tecnologias (e.g. o sinal <TREM> faz referência à roda de trens antigos do tipo Maria-Fumaça, que são cada vez mais raros). Outra possibilidade ocorre quando, uma vez lexicalizados e adequados às restrições fonotáticas da língua, os sinais ganham vida própria e passam a ser submetidos a processos fonológicos e morfológicos que podem vir a alterar a sua estrutura, a ponto de, num dado momento, tornarem-se totalmente opacos no que diz respeito a sua iconicidade.

análise emergem quando se tenta descrever sistematicamente os locais para onde os sinais são direcionados (Liddell, 1990). Segundo Liddell, cada verbo desse tipo carrega, como parte de sua especificação lexical, uma determinada altura em relação ao corpo do falante de onde ele deve partir, e uma determinada altura no corpo do interlocutor para onde deve apontar. Em situações enunciativas concretas, portanto, cada sinal deverá ser apontado de maneira gradiente no espaço de acordo com a altura dos interlocutores presentes – ou mesmo dos interlocutores que, embora ausentes, estejam representados no espaço imediato.

Tomemos como exemplo o sinal <[OBEDECER](#)> da libras, cujos pontos inicial e final envolvem a região da testa.³⁰ Contextualizado numa fala do tipo “Você precisa me obedecer” (i.e. OBEDECER^{X→Y} PRECISAR, considerando-se “x” o referente sujeito e “y” o referente objeto), esse verbo irá assumir diferentes pontos de articulação no espaço dependendo, por exemplo, de os referentes correspondentes a “x” e “y” serem da mesma altura ou não, estarem de pé ou sentados, etc. Cada situação enunciativa particular irá determinar, de acordo com as características do ambiente imediato, a forma que o verbo irá assumir. Para Liddell, o uso de tais verbos envolve, portanto, uma dimensão gramatical convencional, discreta e regrada (relativa à configuração de mão, ponto de articulação, movimento e orientação dos verbos), que é modificada por uma dimensão gestual, gradiente e ad-hoc, que se mostra intimamente atrelada à cada situação enunciativa particular – de uma maneira similar à que ocorre com pronomes como “ele” acompanhados de gesto de apontamento no português. Tal análise, ao ressaltar a importância do gesto na veiculação de sentidos na libras, desloca o fenômeno relativo a esses tipos de verbos do campo morfossintático para o campo da dêixis de pessoa.

Se os verbos indicadores incorporam a gestualidade para expressar referentes do discurso relacionados à ação verbal, os verbos descritivos (*depictive*) incorporam a gestualidade para descrever, ou retratar, certos aspectos de um estado ou ação. Tradicionalmente, tais verbos também receberam denominações diversas, como “verbos de movimento e localização”, “predicados classificadores”, “verbos classificadores”, “verbos polissintéticos”, entre outros. A figura 7 abaixo traz um <[exemplo típico de](#)

³⁰ Para uma descrição inicial das características dos verbos indicadores na libras de acordo com a proposta de Liddell, ver Moreira (2007).

[verbo descritivo na libras](#)>. A narrativa trata de um agricultor que trabalha no campo colhendo pêras em cima duma árvore e, na passagem representada pelas figuras, o agricultor está descendo da árvore para colocar as pêras colhidas num cesto. A mão esquerda do sinalizador representa a árvore, e a mão direita, que se movimenta de cima para baixo a partir do contato com a “árvore”, representa o agricultor.

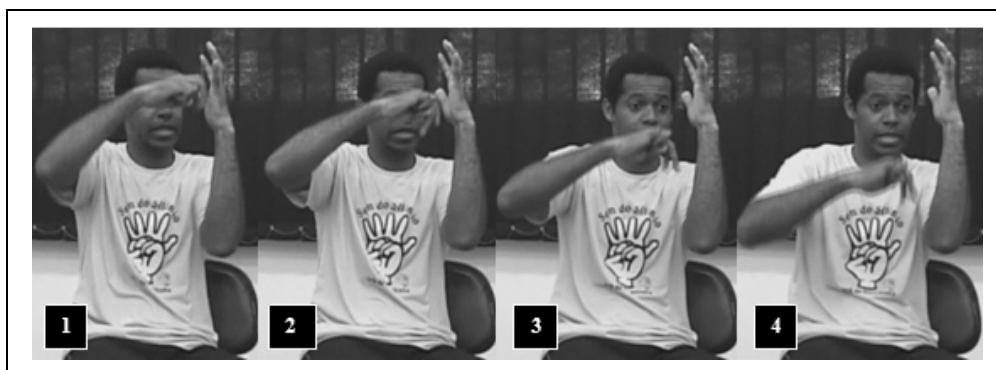


Figura 7. Exemplo típico de um verbo descritivo que representa movimento e localização na libras.

Sob a perspectiva lingüística majoritária no campo das LSs, tais verbos são considerados multimorfêmicos, com cada um dos aspectos que entram na sua composição adquirindo estatuto de morfemas classificadores. A configuração de mão em “V2”, da mão direita, seria um morfema para “seres animados”, como humanos e animais; a configuração de mão em “5”, da mão esquerda, um morfema para objetos compridos verticais, como árvores;³¹ o movimento de cima para baixo, um morfema para “movimento descendente”; o movimento interno de alternância dos dedos – que pode ser mais bem observado no vídeo –, um morfema para o tipo de movimento ao descer da árvore; os pontos de articulação inicial e final da mão direita (quadro 1 e 4), outros morfemas para o local inicial e final da ação realizada; e assim por diante. A expressão sinalizada, realizada por um único sinal, se traduzida em seu sentido literal numa língua oral como o português, envolveria uma longa sentença (e.g. O homem desceu da copa da árvore até o chão com movimentos alternados das pernas), o que caracterizaria um processo de condensação morfêmica similar ao que ocorre em línguas polissintéticas como o esquimó.

Mais recentemente, porém, tal perspectiva sobre expressões desse tipo tem sido debatida. Talmy (2003: 194, nota de rodapé 4), por exemplo, afirma que o principal

³¹ Configuração de mão “V2”:  e configuração de mão em “5”:  .

equivoco está em nomear a expressão toda (i.e. expressões ou predicados classificadores) a partir de um de seus elementos, a configuração de mão, que de fato parece trazer propriedades similares às de classificadores. Estudos como os de Schembri et al. (2005) ajudam a corroborar essa visão. Os autores aplicaram uma bateria de testes – contendo filmes animados que envolviam diferentes tipos de referentes, movimentos e localizações – elaborada para eliciar verbos dessa natureza (Supalla et al., s.d). O mesmo teste foi aplicado a falantes de língua de sinais taiwanesa (TSL), australiana (Auslan) e ouvintes que não falavam nenhuma língua de sinais. Os autores apontam que as descrições dos falantes se mostraram específicas de cada língua apenas no que diz respeito às configurações de mãos, utilizadas de maneira consistente pelos falantes de cada língua de sinais de acordo com as suas convenções particulares. Essas mesmas configurações, no caso dos ouvintes, não apresentaram sistematicidade, o que sugere que, para os ouvintes, elas envolviam a realização de gestos idiossincráticos e ad-hoc e não um conhecimento lingüístico propriamente dito. A retratação de movimento e localização, porém, apresentou um alto nível de similaridade (de 70% a 90%) entre os falantes, independentemente de eles serem falantes de alguma língua de sinais ou ouvintes.

Isso mostra, como argumenta Liddell (2003b), que a representação semântica de certos tipos de entidade nas LSs é realizada por meio de um repertório lexical restrito e convencionalizado de configurações de mãos, algumas delas bem pouco icônicas, diferentemente do que ocorre com a expressão de movimentos e localizações espaciais, que faz uso de recursos gestuais imagéticos, gradientes e ad-hoc. Uma outra evidência para essa proposta é a de que esse tipo de construção da fala, com diferentes camadas de significação envolvendo simultaneamente aspectos lingüísticos categóricos e gestos gradientes, é um fato já documentado nas LOs (e.g. os vários estudos de Adam Kendon, David McNeill e colegas).³²

³² Em um estudo comparativo entre falantes de língua de sinais taiwanesa e inglês a partir de uma narrativa elaborada especificamente para eliciar gestos, Duncan (2005) mostra que essa sobreposição de elementos categóricos e gradientes se reflete também na modificação dos próprios morfemas representados pela configuração de mão. Na narrativa eliciada, um contraste entre duas cenas-chave, uma em que um gato escala uma chaminé por fora, e outra por dentro, num espaço bastante apertado, se reflete na produção dos surdos, entre outras coisas, por meio de “deformações” na forma das configurações de mão convencionais com fins de representar iconicamente a dificuldade de passagem do gato pela chaminé.

Diante da posição estrita de autores como Supalla (2003) – para quem as supostas modificações ad-

Por fim, além da fonologia, morfologia e sintaxe, uma outra área de investigação das LSs que deve ser afetada pelos estudos da gestualidade é a prosódia. Nas próprias LOs – embora essa não seja a visão majoritária sobre a função da prosódia – os trabalhos de Bolinger (e.g. 1961; 1972; 1982; 1983; 1998) sempre tiveram como norte a tese de que a entoação esteja fundamentalmente associada à expressão da emoção, sendo “altamente icônica e deve[ndo] ser estudada em relação a todo o contexto gestual, especialmente as expressões faciais e a linguagem corporal expressiva” (1998: 45). A dimensão gestual da prosódia se revelaria, por exemplo, no fato de os términos de curvas entoacionais serem iconicamente metafóricos e gradientes, de tal modo que a queda no pitch represente encerramento, a subida, continuidade, e quanto maior a queda ou maior a subida, maior o grau de encerramento ou de continuidade (p. 48).

Propostas dessa natureza ganham força com estudos que revelam a coordenação temporal fina entre gestos que co-ocorrem com a fala e a salientação de informações, além da exploração do pitch e do alongamento de sílabas de maneira gradiente para representar iconicamente aspectos de uma imagem (cf. estudos citados em Duncan, 2005). Para McNeill (1992), todo enunciado no uso face-a-face da língua é inevitavelmente formado por unidades lingüísticas categóricas e imagens gradientes, e parece plausível propor que tais imagens sejam expressas por alguns aspectos formais tradicionalmente tratados no âmbito dos estudos da prosódia e da gestualidade.

A discussão do capítulo 4 desta tese, que trata do paralelo entre a função evolutiva do gesto, de um lado, e a função básica de estabelecimento de relações de figura-e-fundo da prosódia, de outro, coloca-se também como uma evidência adicional da relação íntima entre esses dois domínios. Nesse aspecto, a meu ver, a investigação de línguas de modalidade gestual-visual como a libras pode ser reveladora. Tendo em vista o funcionamento dessas línguas num espaço enunciativo essencialmente visual e

hoc de sinais refletem na verdade análises incompletas, não representativas, ou sistemas de transcrição incapazes de captar sistematicidades na forma de contrastes morfêmicos categóricos –, Duncan (2005) destaca, então, a total inversão de valores promovida pelos lingüistas das LSs defensores dessa perspectiva:

Supalla’s comments represent a school of thought in sign language linguistics in which, curiously, it seems that the perspective that gesture is not a part of language is handled, not by distinguishing sign language’s gradient-gestural phenomena from its categorical-linguistic phenomena, so as to exclude gradient phenomena in a principled way from analyses of categorically patterned phenomena, but rather, by a priori exclusion of even the possibility that gradient phenomena are part of everyday sign language use” (p. 311).

espacial, entidades lingüísticas da enunciação podem ser orientadas pelos falantes da mesma maneira com que nós, humanos, dentre várias outras espécies, nos orientamos às entidades do mundo físico para nos socializar e para garantir as nossas necessidades de sobrevivência. Como a análise do capítulo 7 irá mostrar, por meio de recursos gestuais como o direcionamento do olhar, da cabeça, a postura do corpo, gestos de apontamento e a espacialização de sinais, falantes de LSs podem tratar objetos lingüísticos (sejam eles palavras, sintagmas, orações ou mesmo trechos maiores de discurso) como objetos do mundo, salientando certos elementos em detrimento de outros – uma função que nas LOs se mostra tipicamente associada à prosódia.

Se os estudos da prosódia nas LSs são recentes, estudos que analisam a prosódia dessas línguas considerando questões relativas à forma e função da gestualidade são praticamente inexistentes (para uma exceção, cf. Duncan, 2005). Trata-se de uma proposta que ainda carece de investigação, mas que pode trazer um olhar renovado sobre esse domínio da mesma maneira que foi apontado para outros níveis de análise como os da fonologia, a morfologia e a sintaxe.

2.7. Conclusão

O fato de as LSs serem produzidas pelo canal gestual-visual sempre colocou em questão em que medida seria possível estabelecer limites claros entre elementos da gramática (considerados regrados, arbitrários e discretos) e elementos da gestualidade (considerados *ad-hoc*, motivados e gradientes). Pesquisas acumuladas nas últimas duas ou três décadas têm indicado que, embora seja possível separar analiticamente esses dois níveis de produção da língua, a diferença entre ambos parece ser mais de grau do que categórica.³³

As seções anteriores, em especial aquelas referente à morfologia e sintaxe, poderiam ter um tratamento consideravelmente distinto nesta tese, caso a perspectiva aqui adotada envolvesse pressupostos amplamente aceitos por pesquisadores da área, como o da separação categórica entre língua e discurso. Vários dos fenômenos que fonólogos e sintaticistas incluiriam sob o escopo da fonologia, morfologia e sintaxe,

³³ Para duas evidências robustas dessa gradiência nos estudos das LSs, ver Wilcox (2004), que recupera duas rotas por meio das quais gestos cotidianos se tornam parte do inventário lexical e gramatical em diversas LSs; e Taub (2000), que recupera, no pólo da gestualidade, as fontes imagéticas de significação dos verbos descritivos (os chamados “classificadores”), e, no pólo da língua, os caminhos de lexicalização desses verbos em itens com significante e significado altamente convencionalizados.

optei por tratar na seção sobre gestualidade. A razão dessa escolha é a de que vários dos fenômenos que por muito tempo permaneceram obscuros e mal explicados sob uma visão de língua-estrito-senso têm, hoje, diante de uma consideração séria da gestualidade, recebido explicações mais plausíveis – embora ainda haja muito a ser investigado.

Os vários achados apresentados neste capítulo foram trazidos com o intuito, primeiramente, de oferecer um panorama sobre as pesquisas das LSs, mostrando de que maneira as suas gramáticas se organizam em comparação com as das LOs. Além disso, como a discussão acima sugere, buscou-se trazer uma perspectiva sobre esses achados que se mostrasse compatível com o olhar que o presente estudo adota, não apenas sobre a libras mas sobre as línguas humanas de maneira geral. O papel da gestualidade, afinal de contas, não está restrito ao fenômeno das LSs.

3. Fundamentação teórica I: Língua e discurso na perspectiva da análise da conversa

3.1. Introdução

A análise da conversa de base etnometodológica (AC) teve uma importante influência no modo como muitos lingüistas enxergavam a língua, em especial aqueles com orientação ao discurso. O caráter fortemente empírico da abordagem privilegiava uma observação meticulosa de dados naturais ao invés do apelo às intuições de falantes e sentenças descontextualizadas. O interesse pelo modo de pensar do senso-comum exigia um olhar para conversas espontâneas cotidianas ao invés do apelo a gêneros escritos elaborados. O conceito de reflexividade implicava a idéia de que o sistema lingüístico não deve estar desvinculado das ações sociais,³⁴ sendo, portanto, necessário entender como as diferentes ações se estruturam a fim de se entender como a própria língua se estrutura.

No caso da conversação, considerada uma forma primordial de interação social entre humanos, a relação íntima entre a estrutura da prática e a estrutura da língua pode ser entendida, pelo menos em parte, através do fenômeno da projetabilidade. A prática da conversação, enquanto atividade social, exige a coordenação de ações a fim de que certos objetivos sejam alcançados, num processo similar ao que ocorre, por exemplo, na condução de um dueto de música ou dança, ou num cumprimento de mãos (Clark, 1996). Essa coordenação se dá em grande medida pela capacidade que os então-receptores de uma fala têm de identificar as ações num dado turno-em-curso e antever os seus possíveis pontos de completude, alinhando temporalmente suas próprias ações

³⁴ O termo *ação social* será aqui tratado no sentido de Weber (1978: 4-23, apud Schegloff, 2002: 287-288), para quem: “Action is ‘social’ insofar as its subjective meaning *takes account of the behavior of others and is thereby oriented in its course*” (ênfase minha). Utilizando como exemplo uma situação em que dois ciclistas se aproximam em sentidos perpendiculares, sem visão um do outro, Weber argumenta que “a mere collision of two cyclists may be compared to a natural event. On the other hand, their attempt to avoid hitting each other, or whatever insults, blows, or friendly discussion might follow the collision, would constitute social action”.

Embora tal definição se mostre perfeitamente compatível com o empreendimento da AC, Schegloff aponta que essa preocupação – relativa à ação social em seu nível mais básico, da interação direta entre membros – acaba se revelando uma questão marginal no trabalho de Weber, cedendo lugar à discussão de questões tais como a burocracia, a patriarquia, visões de mundo religiosas, etc. Considerando o importante papel da tecnologia na viabilização do estudo de interações cotidianas, então, Schegloff argumenta que um dos intuits da AC é precisamente o de resgatar *essa orientação e esse entendimento* que os participantes revelam uns em relação aos outros, através do registro e análise da forma mais primordial de interação direta entre humanos: a conversação.

às do seu interlocutor de acordo com essas projeções. Essa capacidade dos interlocutores está vinculada à propriedade da fala de apresentar uma estrutura, uma organização, uma ordenação suficiente para permitir tais projeções. O campo da AC, a partir do estudo de Sacks, Schegloff e Jefferson (2003 [1974], doravante SSJ), tem avançado significativamente na análise de fenômenos como esse, revelando padrões de conduta social num nível micro-analítico.

Contudo, o empreendimento da AC sempre foi, antes de tudo, um empreendimento sociológico, não lingüístico. Isso significa que, além da diferença nos objetivos de análise dos sociólogos, certos conceitos fundamentais que se mostravam intimamente relacionados ao uso da língua desenvolvidos nos trabalhos seminais da área permaneceram um tanto vagos e intuitivos. É o caso da chamada “unidade de construção de turnos” (*turn-constructional unit*, ou TCU), que, embora destacada pelos estudiosos da conversação como distinta das unidades lingüísticas por definição, me parecem vez ou outra utilizadas de maneira intercambiável com termos tais como “enunciado” (*utterance*) e “sentença” (*sentence*).

Talvez o problema essencial acerca da noção de TCU seja o fato de que, a despeito de sua definição funcional em termos interacionais (i.e. uma unidade que, ao seu término, constitui um possível ponto de relevância para transição entre falantes), a sua caracterização formal é feita lingüisticamente (i.e. a língua é o principal recurso da fala-em-interação e são os recursos lingüísticos, fundamentalmente, que permitem a antecipação de possíveis pontos de relevância para transição). No que diz respeito a esse pólo formal, desde o início a sintaxe tem sido favorecida pela maioria dos sociólogos na caracterização das TCUs em detrimento de outras dimensões lingüísticas importantes, como a prosódia e a gestualidade. É plausível especular que esse seja um forte viés das transcrições que servem de base para análise, uma vez que a escrita apresenta a língua em uso de forma descorporalizada. Seja como for, esse viés é um dos aspectos que muitos lingüistas trabalhando sob a orientação teórica da AC têm criticado nos últimos anos, destacando, em especial, a importante contribuição da prosódia para a organização da conversação.

Neste capítulo, então, discutirei em maiores detalhes a visão da AC sobre a relação entre língua e discurso (i.e. gramática e interação), visão essa que tem servido como base para a abordagem da atual pesquisa. Especificamente, pretendo introduzir algumas pesquisas representativas que, direta ou indiretamente, trataram do fenômeno

da projetabilidade; e articular uma proposta de caracterização dos conceitos de *turno*, *TCU* e *ponto de relevância para transição* que se mostre psicológica e interacionalmente plausível, tendo em vista a importância desses conceitos para o nosso entendimento sobre o trabalho de coordenação da participação na conversação.

Uma das preocupações centrais do capítulo é a de revelar o forte viés em favor dos recursos léxico-sintáticos para o trabalho de projeção da fala em alguns estudos clássicos da AC e, alternativamente, recuperar a importância da gestualidade e principalmente da prosódia nesse sentido. Considerando-se o problema da segmentação gramatical da libras, que ocupa a presente pesquisa, a conclusão geral é a de que, embora a AC ofereça um referencial teórico fundamental para a abordagem dessa questão, oferecendo uma perspectiva sobre a relação entre estrutura e contingência na fala espontânea, a sua definição sobre a forma e função das unidades mínimas do discurso ainda se mostra limitada. Ainda que estudos lingüísticos provenientes da lingüística funcional e da lingüística cognitiva ofereçam uma alternativa importante nesse aspecto da análise, esse olhar alternativo – a ser articulado no capítulo 4 – deverá ser trazido neste capítulo apenas tangencialmente, isto é, na medida em que possibilite apontar os pontos da abordagem da AC que, a meu ver, carecem de reformulação.

3.2. A segmentação da língua e a troca de turnos

A proposta de iniciar o estudo de segmentação gramatical a partir da interação face-a-face pode parecer uma proposta pouco ou nada apropriada para lingüistas que fazem uma cisão categórica entre língua/fala ou competência/performance, atribuindo à língua em uso uma natureza imprevisível e muitas vezes caótica. Apesar disso, especialmente após o surgimento dos gravadores de áudio e vídeo, um número cada vez maior de pesquisas tem oferecido evidências de que a língua na interação espontânea é não apenas passível de uma análise rigorosa, mas também de grande interesse teórico. Estudos com tal orientação têm revelado o fato de que a conversação, ao contrário do que se pensava, não é desordenada, mas apresenta uma grande sistematicidade e deve estar intimamente relacionada ao nosso conhecimento abstrato da língua.

Sintomaticamente, parte substancial das pesquisas que impulsionaram a descrição da língua na interação partiu do trabalho de estudiosos de outras ciências humanas, como a sociologia (e.g. Goffman, 1967) e a antropologia (Gumperz e Hymes, 1972), uma disciplina que já há muito tempo influenciava os rumos da lingüística

americana (Sapir, 1921). Entre esses trabalhos, destaca-se o estudo sociológico seminal de SSJ sobre o sistema de tomada de turnos na conversa espontânea e os demais trabalhos que têm constituído o campo da AC.³⁵ Desde esse estudo, o olhar sociológico e empirista sobre a gramática tem se mostrado uma proposta alternativa extremamente produtiva para o estudo das línguas naturais, o que se evidencia pelo número cada vez maior de trabalhos lingüísticos no campo recentemente denominado *lingüística interacional* (Selting e Couper-Kuhlen, 2001).

Vista a partir do olhar etnometodológico que a fundamenta, essa vertente trata a gramática como uma estrutura formal altamente regulada, embora sensível ao contexto, que está disponível como recurso e que é imposta como restrição para a realização de ações práticas pelos membros de um grupo em sua vida social cotidiana (Garfinkel e Sacks, 1970). É o que fica claro nas palavras de Schegloff (1996) sobre a abordagem geral para o estudo da gramática na interação:

The central prospect ... is that grammar stands in a reflexive relationship to the organization of a spate of talk as a turn. On the one hand, the organizational contingencies of talking in a turn ... shape grammar – both grammar as an abstract, formal organization and the grammar of a particular utterance. On the other hand, the progressive grammatical realization of a spate of talk on a particular occasion can shape the exigencies of the turn as a unit of interactional participation on that occasion, and

³⁵ No Brasil, trabalhos voltados para o campo denominado “análise da conversação” ou mesmo para o estudo da língua oral que surgiram na década de 80 parecem perder – ou não compartilhar, como deixa claro Marcuschi (1986: 21) – a perspectiva etnometodológica da linha de pesquisa conhecida como *conversation analysis*, originada nos EUA.

Metodologicamente, os trabalhos brasileiros divergem dessa tradição americana, ora pela diferença marcante das convenções e do nível de detalhamento das transcrições, ora pela investigação de outros tipos de fala-em-interação (e.g. entrevistas) em termos de conceitos desenvolvidos para a conversação, sem, a meu ver, a devida consideração das diferenças que esses gêneros acarretam.

Analicamente, essa divergência se manifesta pelo menos de duas maneiras. Em primeiro lugar, certos conceitos perdem uma dimensão crucial de sua definição, a saber, o fato de serem sensíveis ao contexto, às contingências da conversação em tempo real – característica que, como ficará claro na exposição deste capítulo, marca o fenômeno da projetabilidade; em segundo lugar, o propósito de análise parece ser muito mais o de desenvolver tipologias estanques (e.g. tipos de turnos, tipos de “correções”, etc) que possam ser “aplicadas” à análise de diferentes instâncias de uso oral da língua, do que o de analisar as estratégias engenhosas dos participantes, planejadas e levadas a cabo em tempo real, para lidar com as contingências da interação em face de seus propósitos práticos – o modo de pensar do senso-comum, que constitui o interesse central da etnometodologia.

Acredito que o meu trabalho esteja teórica e metodologicamente alinhado a trabalhos brasileiros mais recentes, desenvolvidos sob a área denominada “análise da conversa de base etnometodológica” (e.g. Gago, 2004; Garcez, 2001; Garcez e Loder, 2005; Loder et al., 2004; ver também Ribeiro e Garcez, 2002), que, como o próprio nome demonstra, busca justamente recuperar o vínculo teórico etnometodológico que a vertente brasileira da “análise da conversação” deixou de lado.

the grammatical properties of a language may contribute to the organization of turns-at-talk in that language and of the turn-taking device by which they are deployed (p. 56).

Considerando essa relação reflexiva proposta na AC entre a organização gramatical das línguas e a organização da conversação – relação também sugerida no estudo de Baker e Padden (1978), introduzido no capítulo 1 – a presente pesquisa buscou verificar de que maneira os conceitos operacionais e o conhecimento acumulado no âmbito da AC poderiam contribuir para elucidar a questão da segmentação da libras a partir da análise de conversas entre surdos adultos fluentes.

3.3. A proposta original de SSJ

O trabalho seminal de SSJ propõe uma sistemática para o processo de troca de turnos na conversa espontânea que, pela elegância teórica e forte sustentação empírica, teve uma grande repercussão, não apenas nos estudos sociológicos mas também nos estudos lingüísticos. Embora a questão da tomada de turnos já tivesse sido abordada por diversos autores em diversas linhas de pesquisa, SSJ argumentam que seu mecanismo de organização e operação quase nunca era analisado como um fim em si, mas sempre como *meio* para outras análises que tomavam por base essa organização. Por esse motivo, os autores deslocam o foco de interesse para o entendimento desse mecanismo específico, buscando elaborar um aparato formal que possa dar conta de uma série de constatações empíricas levantadas a partir da observação e análise de conversas em um corpus de língua inglesa.

Essas constatações empíricas podem ser resumidas da seguinte maneira. Numa conversação, predominantemente um fala de cada vez. Apesar disso, há troca de falantes, há variação do tamanho dos turnos, e há variação na ordenação dos falantes. Além disso, as transições de um falante para o outro são coordenadas de maneira fina, havendo poucas sobreposições de falas ou intervalos entre as falas, em muitos casos a transição ocorrendo sem qualquer sobreposição ou intervalo. Tais observações levaram os autores a crer que deva haver técnicas para alocação de turnos e para a construção de enunciados de tal modo que essas técnicas permitam a coordenação das transferências de fala e alocação dos falantes.

A elegância teórica do sistema elaborado por SSJ reside na sua natureza ao mesmo tempo livre de contexto e sensível ao contexto (p. 13-14). Por um lado, a necessidade de o sistema ser livre de contexto está relacionada ao fato de a conversa

espontânea: a) poder acomodar uma grande variedade de situações; b) ser um canal de interação por meio do qual pessoas de diferentes identidades e diferentes grupos continuamente operam; e c) ser capaz de acomodar mudanças de situação dentro de uma mesma conversa. Por outro lado, a necessidade de o sistema ser sensível ao contexto provém do fato de a ordenação dos turnos variar, o tamanho dos turnos variar, a extensão das conversas variar, a distribuição dos turnos variar, entre outros fatos que evidenciam a sensibilidade do sistema às contingências do contexto imediato.

Com base nessas considerações, SSJ propõem a existência de dois componentes do mecanismo de organização da tomada de turnos, acompanhados de algumas regras.³⁶ O primeiro é o chamado componente de construção de turnos (SSJ: 16), de acordo com o qual os falantes dispõem de vários tipos de unidades, desde construções lexicais até sentenciais, para a construção dos turnos na fala. A ocorrência de um dado tipo num turno específico permite uma projeção aproximada por parte do interlocutor do tipo de unidade que está em curso e, conseqüentemente, do ponto no turno onde essa unidade poderá estar completa. O falante tem o direito de prosseguir o seu turno sem interrupção até o término da unidade em curso e o primeiro *possível* ponto de término da unidade é chamado de “ponto de relevância para transição” (*transitional-relevance place*, ou TRP). A transferência da condição de falante entre os interlocutores é coordenada com base nesses TRPs.

O segundo componente refere-se ao processo de alocação de turnos (SSJ: 16-17), de acordo com o qual duas técnicas básicas podem ser identificadas: aquelas em que o turno seguinte é alocado pelo falante corrente, que seleciona o falante seguinte (e.g. por meio do emprego da primeira parte de pares adjacentes tais como chamado-resposta, elogio-aceitação/rejeição, etc); e aquelas em que o turno seguinte é alocado por auto-seleção, quando algum dos interlocutores opta por dar continuidade à sua participação na interação sem que tenha sido selecionado pelo falante-corrente.

Operando sobre esses componentes, as regras a seguir permitem a alocação do turno seguinte a um único participante (i.e. o princípio do um-de-cada-vez) e coordenam

³⁶ O uso do termo “regras” no trabalho original de SSJ gerou controvérsias e interpretações equivocadas a respeito da natureza do sistema (Searle, 1992). Assim, cabe destacar, como argumenta Schegloff (1992), que as regras descritas devem ser entendidas como princípios frente aos quais os falantes se orientam de maneira tácita, e não, por exemplo, como análogas a regras de trânsito, como Searle sugeriu.

a transferência de modo a minimizar os intervalos e as sobreposições de fala. De acordo com essas regras, em qualquer turno:

(1) Para qualquer turno, no primeiro lugar relevante para a transição de uma primeira unidade de construção de turno:

(a) Se o turno até aqui está construído de modo a envolver o uso de uma técnica de ‘falante corrente seleciona o próximo’, então a parte assim selecionada tem o direito e é obrigada a tomar o turno seguinte para falar; nenhuma outra parte possui tais direitos ou obrigações, e a transferência ocorre naquele lugar.

(b) Se o turno até aqui está construído de modo a não envolver o uso da técnica de ‘falante corrente seleciona o próximo’, então a autoseleção para a próxima vez de falar pode ser instituída, mas não necessariamente; quem inicia primeiro adquire o direito ao turno, e a transferência ocorre naquele lugar.

(c) Se o turno até então é construído de forma a não envolver o uso da técnica de ‘falante corrente seleciona o próximo’, então o falante corrente pode, mas não precisa continuar, a menos que outro se auto-selecione.

(2) Se, no primeiro lugar relevante para a transição de uma primeira unidade de construção de turno, nem 1a e nem 1b operaram, e, seguindo a provisão 1c, o falante corrente continuou, então o conjunto de regras a–c reaplica-se no próximo lugar relevante para a transição e recursivamente a seguir em cada lugar relevante para a transição, até a transferência ser efetivada (p. 16-17).

A questão que mais interessa à presente discussão é o papel da gramática nesse modelo. Como argumentam SSJ, “[s]e examinamos materiais empíricos para ver onde, em um turno em andamento, os próximos falantes começam (ou tentam começar) turnos seguintes, percebe-se que tais começos não ocorrem continuamente ao longo do curso em desenvolvimento de um turno, mas em pontos específicos no decorrer do seu desenvolvimento” (p. 41). De modo geral, a localização desses possíveis pontos discretos no turno coincide, segundo os autores, com os possíveis pontos finais de unidades sintáticas, ainda que eles chamem a atenção para a necessidade de considerar o papel da prosódia como recurso estratégico de modificação desses pontos.

Ao fazer uso de um termo como *unidade de construção de turnos*, ou TCU, contudo, Schegloff argumenta que a questão não é simplesmente a de propor uma mudança terminológica – de sentenças ou orações para TCUs (Schegloff, 1996: 55-6). Mais do que isso, a idéia é a de que a fala-em-curso na conversação se mostre aberta a “considerações relativas à reatividade e ao valor interacional, ao ajuste ao interlocutor, à

recalibragem momento-a-momento, à re-organização e ao re-acabamento, e à co-construção interacional”. O fato de a gramática que habita turnos-na-fala ser realizada em tempo real para os participantes na conversação leva o autor, então, a afirmar que “se ‘sentenças’, ‘orações’ e ‘sintagmas’ vão acabar se mostrando implicados [na conversação], eles serão diferentes em ênfase, e talvez em tipo, dos objetos sintáticos estáticos de grande parte da teorização lingüística”.

A princípio, poderia ser argumentado que a discrepância entre os objetos estáticos da língua e a dinamicidade da fala resulta precisamente da distinção entre competência e performance, ou língua e fala, e que o apelo dos analistas da conversação em favor de uma gramática dinâmica e flexível, portanto, não necessite ser mantido. Contudo, diferentemente dos pressupostos clássicos da lingüística, a AC tem oferecido amplas evidências de que as instâncias concretas da fala, inclusive as chamadas “disfluências”, não devem ser assumidas como falhas na implementação de um sistema perfeito sob determinadas restrições cognitivas e emocionais. Ao contrário, elas são recursos metódicos convencionais levados a cabo pelos participantes estrategicamente a fim de realizar trabalhos interacionais específicos (e.g. resolução de instâncias de sobreposição de fala (Schegloff, 2000), sinalização de elementos reparáveis problemáticos e seus reparadores (Schegloff, 1979), sincronização da atenção dos interlocutores sobre o que é dito (Clark, 2002), entre outras propostas).

Essa reflexão revela que o olhar para a gramática sob a ótica da AC exige uma teorização de língua que leve em consideração a sensibilidade dos enunciados ao processo contingente e colaborativo de uso da língua produzida em tempo real. O conceito de TCU parece ter sido elaborado exatamente em vista da incompatibilidade atestada pelos analistas da conversação entre as unidades observadas no uso concreto da língua, com toda sua dinamicidade e flexibilidade, e as unidades gramaticais tradicionais da lingüística, altamente abstratas e estáticas.

Por outro lado, já na proposta inicial de SSJ nota-se um viés analítico em favor da sintaxe como elemento central da gramática – e, interacionalmente, como elemento central para o estabelecimento da coordenação da ação social – que me parece muito próximo do viés que os próprios lingüistas têm trazido ao longo dos tempos. Esse favorecimento da sintaxe, explicitamente reconhecido por Schegloff em trabalho subsequente (1998: 237, nota 3), tem sido revisto e/ou criticado por pesquisadores, principalmente lingüistas, que, dando prosseguimento à reflexão sobre como a projeção

de TRPs é realizada na conversação espontânea, têm proposto redefinições e aprofundamentos sobre o conceito de TCU. Algumas dessas propostas serão discutidas a seguir.

3.4. A projetabilidade dos turnos-na-fala

Ao reconhecer o papel dos recursos lingüísticos para a projeção de pontos de transição entre falantes e, conseqüentemente, para o estabelecimento da coordenação de ações sociais na interação, SSJ destacam, já em seu artigo seminal, que essa é uma área para a qual os lingüistas poderiam dar grandes contribuições. De fato, desde então, o número de trabalhos lingüísticos voltados para a inter-relação entre gramática e interação, em especial no que concerne à dimensão operacional da organização da troca de turnos, tem crescido significativamente com o passar dos anos.

Uma das principais contribuições desses estudos para a compreensão acerca do fenômeno da projetabilidade da conversação cotidiana foi a identificação de outros recursos lingüísticos, além dos lexicais e sintáticos já apontados por SSJ, operando no trabalho de coordenação de ações sociais. Entre esses recursos, destacam-se os fonéticos (e.g. Local e Kelly, 1986; Local e Walker, 2004) e prosódicos (e.g. vários trabalhos em Couper-Kuhlen e Selting, 1996; Ford e Thompson, 1996), os gestuais (e.g. Goodwin, 1981; Streeck e Kallmeyer, 2001),³⁷ além dos próprios recursos sintáticos e seqüenciais que continuam a ser investigados considerando-se especificidades de língua para língua (e.g. Auer, 2005; vários trabalhos em Ochs et al., 1996; Tanaka, 2000). Ademais, recursos extralingüísticos, tal como a exploração do ambiente físico e cultural imediato, também foram considerados e têm se mostrado relevantes para esse trabalho de coordenação da ação social (e.g. Clark, 2005; Mondada, 2006).

Alguns dos estudos nessa área de investigação exigem aqui uma consideração mais detalhada por introduzir propostas que, a meu ver, contribuem de maneira significativa para o aprofundamento da discussão. Um desses estudos é o de Houtkoop e Mazeland (1985), que analisaram o trabalho dos participantes para a realização de atividades discursivas maiores, como a narração de histórias. Cunhando o termo *unidades discursivas* para se referir às unidades de diferentes tipos e tamanhos da fala-

³⁷ Neste estudo, a menos que seja especificado, usarei os termos “gesto”, “gestual” e “gestualidade” de maneira ampla, englobando não apenas as gesticulações manuais mas também expressões faciais, orientação e movimentos corporais.

em-interação, os autores estabelecem um paralelo entre unidades do tipo histórias e unidades gramaticais. Em ambas, seria possível identificar uma estrutura envolvendo uma trajetória de início, meio e fim, frente à qual os participantes se orientam para a organização de sua participação. Por exemplo, tão logo uma história (pré)-anunciada tenha sido ratificada pelo receptor, esse irá orientar a sua participação de acordo com o possível ponto de completude da narrativa, projetado com base no prefácio oferecido pelo narrador, num trabalho de projeção que seria similar ao que ocorre com as unidades gramaticais.³⁸

Tendo em vista que a execução de atividades amplas como as histórias envolve um longo processo de construção progressiva, incrementado momento a momento, os autores argumentam que, no contexto de atividades dessa natureza, os TRPs a princípio vinculados às unidades mínimas seriam “bloqueados” até que uma unidade reconhecível como final fosse produzida. Em vista disso, as unidades que constituem uma história seriam caracterizadas por uma incompletude inerente, e diante delas os receptores tenderiam a restringir a sua participação a breves sinais de acompanhamento que não almejam o controle da *palavra*.³⁹ Os autores propõem, então, que o modelo de SSJ seja

³⁸ Um exemplo hipotético de prefácio de uma história no contexto conversacional seria: “Puts, você não sabe a besteira que eu fiz no arquivo da minha tese um dia antes de ter que entregá-la”. Esse tipo de enunciado tem a função, de acordo com Sacks (1992), de solicitar uma resposta do receptor ratificando a realização da narrativa. Uma vez autorizada a história, então, o falante sabe que poderá fazer uso de um turno de fala mais extenso para levar a cabo a sua narrativa e o receptor sabe que deverá orientar a sua participação principalmente em relação ao enunciado que trouxe o clímax da história projetado pelo prefácio (no caso, uma besteira no final de produção de uma tese).

³⁹ Nessa discussão, considerando-se a expressão vernacular de “estar com a palavra”, utilizarei o termo “palavra” como tradução de *floor*. Seguindo essa mesma orientação, seria mais consistente traduzir *turn* por “vez”, mas o termo “turno” já está altamente estabelecido na literatura sobre interação no Brasil.

Entendo a idéia de “ter a palavra” (*having the floor*) como associada à idéia de “controlar por direito um espaço virtual de interação”. No caso em questão, o argumento é o de que as participações dos receptores durante a narração de uma história não almejariam a transferência desse controle, estando limitadas a certas práticas (e.g. pedidos de esclarecimento, sinais de acompanhamento, demonstrações de interesse) que, a despeito de envolverem a tomada de um turno, visam a manter o controle da atividade em curso (no caso, a narrativa) com o falante-corrente.

Apesar dessa relevância analítica da distinção entre *turno* e *palavra*, a sistemática proposta por SSJ trata esses termos sem uma definição muito precisa. Por exemplo, excetuando os casos em que um falante-corrente seleciona o falante-seguinte (e.g. endereçando-lhe uma pergunta), um interlocutor tem três possibilidades de participação diante de um TRP, ao término possível de uma TCU: a) tomar a palavra, fazendo uso de um turno pleno; b) não tomar a palavra, permanecendo em silêncio; ou c) não tomar a palavra, utilizando um tipo de turno que evidencie o seu estatuto primário de receptor. Essa última opção não foi vislumbrada no artigo seminal, que não mostra uma preocupação em distinguir as tomadas de turnos que coincidem com a tomada da palavra, de um lado, e a tomadas de turnos que não envolvem a tomada da palavra, de outro. Desse modo, o argumento de SSJ de que “a transferência do

modificado para dar conta do funcionamento do sistema de tomada de turnos sob a restrição de atividades interacionais específicas.⁴⁰

Como aponta Ford (2004: 33), contudo, é preciso destacar uma limitação no trabalho de Houtkoop e Mazeland, uma vez que seus dados não incluem instâncias de histórias nas quais a participação dos então-receptores *durante* a sua implementação vai além das contribuições minimalistas não-competitivas. É o caso de histórias que acabam atrapalhadas ou mesmo fracassadas em meio à sua realização pelo fato de, por exemplo, o então-receptor iniciar a sua própria história, ou fazer contínuas provocações visando a causar constrangimento e incômodo ao contador, ou colocar em questão a pertinência da narrativa, entre outras possibilidades (e.g. Goodwin, 1986).

Embora a proposta de Houtkoop e Mazeland sobre as TCUs (ou unidades discursivas, em sua terminologia) seja ainda insatisfatória, a sua discussão sobre atividades mais extensas é esclarecedora na medida em que revela uma aparente contradição inerente à definição do conceito de TCU. Se, de fato, os prefácios de história visam a postergar a relevância para a transição do estatuto de falante, autorizando a realização de um turno extenso por parte de um dos interlocutores (como argumenta Sacks, 1992: 227); e se, de fato, essa relevância para transição é definidora do conceito de TCU (como argumenta Schegloff, 1996); então a conclusão lógica é a de que a história como um todo, independente de seu tamanho, constituiria uma única TCU complexa, internamente formada por múltiplas unidades de alguma outra natureza não-especificada (ver o argumento em Selting, 2000: 485-6). Essa extensão lógica da definição de TCU, contudo, se choca com os fatos empíricos, que, a despeito de

estatuto de falante é coordenada em referência a esses pontos de relevância para transição” (p. 12) obscurece o fato de que, nesses mesmos pontos, uma troca de turnos que não coloque os papéis dos interlocutores em questão também seja relevante – essa discussão será retomada ao longo do trabalho, em especial na proposta sobre o conceito de TRP na seção 3.5.

Para uma discussão sobre a indefinição dos termos *turno* e *palavra* na literatura voltada para a interação, ver Edelsky (1981: 397-402).

⁴⁰ Sacks (1992) já discutia o fato de que, na realização de atividades sociais mais extensas como histórias, a relevância para a transição seria deslocada para o final da atividade como um todo, não se sucedendo a cada sentença individualmente. Para ele, a relevância de cada uma das sentenças constitutivas de uma história se manifestaria na sua inter-relação mútua, e não de maneira independente. Contudo, em suas aulas, Sacks freqüentemente discutia seus *insights* utilizando termos como *sentence* e *utterance* e é difícil afirmar com segurança de que modo eles se relacionam especificamente com os conceitos de TCU e TRP (e.g. Sacks, 1992: 226-227). Assim sendo, prefiro assumir que a posição de Sacks esteja mais bem representada no artigo seminal de SSJ, no qual não há referência a modificações da relevância para transição de turnos em contextos de atividades sociais específicas, como Houtkoop e Mazeland propõem.

qualquer tendência mais comum, mostram como a construção progressiva das histórias está, na verdade, continuamente suscetível a incursões do receptor – ainda que tais incursões *tendam* a ser restritas a determinadas práticas. Essa aparente inconsistência será retomada na seção 3.5, em que uma proposta sobre os conceitos em questão será articulada.

Enquanto atividades extensas como histórias são claramente construídas por meio de incrementos sucessivos, o caráter incremental e contingencial de unidades gramaticais é algo pouco consensual. Nesse sentido, Goodwin (1979) discute o processo de produção de um enunciado a fim de demonstrar que a abertura da fala-em-interação à co-construção também opera no nível das unidades gramaticais, e não apenas no de atividades mais extensas como as histórias. Analisando o enunciado produzido por John, “I gave, I gave up smoking cigarettes::. I-uh: one-one week ago t’da:y. acshilly”,⁴¹ o autor busca demonstrar que a realização da unidade abstrata que os lingüistas chamam de sentença está constantemente submetida às contingências da interação, constituindo-se numa construção colaborativa que depende do falante e de seu(s) interlocutor(es) para ser realizada. Para isso, o autor analisa o processo de produção do enunciado passo a passo, mostrando como as mudanças de direção do olhar de John em busca de novos interlocutores, aliadas ao seu nível de conhecimento compartilhado com cada um deles, explicam a forma aparentemente bizarra que a “sentença” assume nesse contexto.

É importante destacar, porém, que a dinamicidade da passagem se torna muito mais marcante quando consideramos, assim como Goodwin o faz, a *sentença* reconstruída de maneira *post-hoc* a partir da passagem, ao invés dos sucessivos incrementos prosódicos *ad-hoc* que de fato constituem a fala de John. A meu ver, o autor superestima o valor da sentença como ponto de partida para análise quando afirma, sobre a passagem, que: “a coerência manifesta de seus enunciados como uma única sentença constitui tanto uma observação inicial sobre a sua organização quanto *uma garantia* para a análise dessa fala como uma única unidade” (p. 98, itálico meu).

Uma alternativa de análise mais “garantida”, contudo, seria a de considerar a passagem como uma sucessão de fragmentos, ou melhor, incrementos, cada um deles

⁴¹ O ponto final (.) indica contorno prosódico final descendente; a vírgula (,) indica contorno final ascendente; os dois pontos (:) indicam alongamento do segmento; os grifos () indicam acento. Além disso, entre as expressões “cigarettes” e “I-uh”, um dos participantes emite um sinal de acompanhamento – “Yeah” –, seguido de 0,4 segundos de pausa, que não foi representado aqui.

formulado e levado a cabo de acordo com as contingências emergentes em tempo real (i.e. o (não) estabelecimento de contato visual de certos interlocutores com o falante e os seus diferentes níveis de relação com ele). Sob essa perspectiva mais fragmentária, a abertura da fala ao re-acabamento e à reorientação gramatical se manifestaria principalmente *entre* cada um dos incrementos – em decorrência da demanda de *formulação*,⁴² em tempo real, de cada novo incremento em face de cada nova contingência encontrada – e não no âmbito de uma construção gramatical altamente abstrata denominada “sentença”. Essa alternativa não invalida, em absoluto, a análise de Goodwin; apenas sugere que as construções gramaticais, em especial as mais complexas (e.g. sentenciais), devam ser um epifenômeno do processo local de formulação da fala-em-interação espontânea; e que, até que uma “sentença” se complete, vários pontos de relevância para transição devam emergir em face dos sucessivos fragmentos que contigencialmente a compõem.⁴³

⁴² O conceito de *formulação* (Sacks, 1992; Schegloff, 1972) será usado neste trabalho em contraste com a idéia de *implementação* de planos pré-fabricados. Embora ambos envolvam um claro investimento cognitivo, a formulação tem uma dimensão social e contingencial intrínseca que, em geral, não aparece associada às idéias de planejamento e implementação, freqüentemente vistas como atos individuais e pré-determinados. A formulação, diferentemente, implica um trabalho de categorização *seletiva*, desenhada para um interlocutor específico, num momento específico de uma situação específica, a fim de atingir determinados propósitos interacionais. Em vista desse caráter seletivo, as formulações dos participantes na fala-em-interação os apresentam como posicionados no mundo e moralmente responsáveis por suas posições adotadas (Garfinkel e Sacks, 1970).

Como exemplo, um mesmo referente X pode ser formulado de modo a sinalizar diferentes aspectos de sua identidade *a partir da perspectiva do formulador*. Uma determinada pessoa pode ser chamada de “ele”, “o João”, “o cara”, “o Jô”, “o simpático vendedor da esquina”, entre outras incontáveis denominações que, uma vez enunciadas, irão situar o formulador como socialmente posicionado no mundo e responsável por essa posição. Além disso, ainda que os sociólogos tenham implicitamente restringido as análises das formulações à (dimensão mais saliente da) escolha lexical (ver, e.g. Schegloff, 2007: 463), tão relevante quanto a escolha de palavras e expressões é a escolha da configuração sintática, prosódica e gestual que inevitavelmente constituem a fala espontânea (ver discussão da seção 4.3 sobre o conceito de *estruturação ideacional*, cf. Langacker, 1987; e os vários trabalhos de Dwight Bolinger sobre os efeitos de sentido resultantes de diferentes combinações de gesto e entoação).

Desse modo, mais do que considerações de ordem cognitiva e discursiva sobre o estatuto de certos referentes (a perspectiva central das análises funcionalistas sobre referência), um aspecto crucial na escolha de uma formulação são os propósitos práticos do produtor diante de uma determinada contingência imediata (e.g. demonstrar uma relação neutra, íntima, distante, antagônica, etc em relação a esse referente). A essência da conversação é menos a troca neutra de conteúdos substantivos do que a negociação de posições e relações sociais (Goffman, 1974: 511-516).

⁴³ A notação e inclusive a análise de Goodwin mostram que há vários pontos de relevância para transição ao longo da dita “sentença”, um deles, em especial, seguido por um sinal de acompanhamento (“Yeah”, após o final de “cigarettes”) e outro que, diferentemente, não é seguido de resposta do interlocutor (após “today”, como argumentado pelo autor na p. 110-111). Todos esses pontos coincidem com os termos de contornos prosódicos, em alguns casos acompanhados de alongamentos finais característicos desses

Em vista da dinamicidade marcante da conversação, que Goodwin demonstra de maneira persuasiva por meio desse estudo de caso, Streek (1995) argumenta que a questão da projeção de unidades na fala-em-interação deva sempre ser entendida como um processo colaborativo dos participantes, o que permite contrastar a idéia de projeção com a de determinação (ver também Auer, 2005: 8-9). Passível de falha, a projeção de uma unidade pode ser rompida ou ter a sua trajetória alterada antes que um dado elemento seqüencial previsto seja alcançado. Para que se alcance o final projetado de uma unidade, então, a participação do receptor é decisiva. Isso torna a projetabilidade um fenômeno não exatamente cognitivo e unilateral (de instanciação de, e inferência sobre, unidades pré-fabricadas), mas inerentemente social e colaborativo (de seleção, negociação, adequação, interpretação e ratificação mútua de unidades relevantes cognitivamente com base nas contingências e propósitos práticos dos participantes da situação imediata) (ver, por exemplo, Ford, 2004).

Consideradas as características gerais do fenômeno da projetabilidade, resta ainda refletir sobre como os recursos lingüísticos se inter-relacionam a fim de estabelecer pontos relevantes para transição, ou TRPs, na cadeia da fala. Da discussão de Goodwin, em especial, já fica evidente o estatuto privilegiado que tenho conferido à segmentação prosódica (em comparação com o favorecimento de SSJ em relação à segmentação sintática) para o estabelecimento desses pontos. Esse favorecimento será retomado em vários pontos deste trabalho e será mais bem justificado no capítulo 4. Ainda no âmbito dos estudos da interação, porém, um trabalho em particular oferece importantes evidências para essa visão: o de Ford e Thompson (1996).

Em seu estudo, as autoras analisam o papel da inter-relação entre recursos sintáticos, prosódicos e pragmáticos para a constituição de TRPs. Estabelecendo critérios para identificação de unidades em cada um dos três níveis, as autoras constatam que, de fato, a convergência do término de unidades desses diferentes níveis, estatisticamente, é o fator preditivo mais robusto para indicar os pontos da conversa em

términos. Isso sugere que o argumento central de Goodwin não difira de fato do que eu aqui proponho, e que a idéia de estudar a contingencialidade “no processo de construção de uma sentença na língua natural” seja uma *formulação* que deve ser entendida em seu caráter situado: o diálogo de Goodwin (1979) com a ciência lingüística de sua época e suas visões acerca das propriedades da língua. Apesar disso, tendo em vista que alguns trabalhos subseqüentes têm conferido um peso a meu ver excessivo ao conceito de “sentença” – como o de Lerner (1996), a ser discutido adiante –, achei importante apontar desde já de que modo o trabalho de Goodwin poderia ser interpretado sob um viés distinto.

que, empiricamente, se opera a transição de falantes (71% das ocorrências). Os pontos no turno que apresentam essa convergência de fatores são chamados por elas de “complexos pontos de relevância para transição” (CTRPs), em contraposição aos TRPs definidos por qualquer um desses fatores isoladamente. A partir dessa análise essencialmente quantitativa, as autoras analisam então, qualitativamente, as instâncias em que a troca de falantes se dá em pontos onde a convergência sintática, prosódica e pragmática *não* ocorre. Essas “violações” da norma, então, são explicadas em termos de práticas sociais já documentadas no campo na AC que, embora envolvam também a orientação dos participantes em relação aos CTRPs, são desenhadas com o intuito específico de violar estrategicamente esses pontos, em vista de determinados fins interacionais.⁴⁴

Embora, à primeira vista, a análise de Ford e Thompson (1996) pareça conferir igual peso à sintaxe e à prosódia, um olhar mais cuidadoso sobre esse trabalho mostra um claro favorecimento da dimensão prosódica. Do ponto de vista teórico, esse favorecimento se revela no fato de que o final de um contorno entoacional é utilizado como critério formal não apenas da possível completude prosódica, mas também da completude pragmática. Embora as autoras ressaltem que isso não implique uma coincidência direta entre ambos os fatores, uma vez que muitos finais de contornos prosódicos podem ou não constituir ações completas no nível pragmático, fica evidente que, para elas, uma ação social não pode ser vista como completa no meio da realização de uma unidade entoacional.⁴⁵ Do ponto de vista empírico, esse favorecimento da prosódia se revela no fato de a organização sintática apresentar um número muito maior de possíveis pontos de completude em relação à organização prosódica e pragmática e, apesar disso, a troca de falantes não corresponder *sistematicamente* a esses pontos. Como as próprias autoras argumentam, a sintaxe, por sua natureza recursiva, oferece

⁴⁴ Por exemplo, numa análise sobre a alocação do início de sobreposições de fala, Jefferson (1973) argumenta que sobreposições prematuras em meio ao turno-em-curso, i.e. em pontos distantes do TRP ou mesmo do espaço de transição, não indicam erros de coordenação temporal ou mesmo falta de orientação em relação às TCUs e TRPs. Diferentemente, tais incursões emergem estrategicamente em *pontos reconhecíveis* ao longo da trajetória da unidade a fim de sinalizar ao falante corrente, por exemplo, que o então-receptor já tem conhecimento sobre o conteúdo do turno “interrompido”.

⁴⁵ Apesar de essa posição teórica parecer digna de uma justificativa, as autoras não a discutem. Tendo em vista que essa me parece a posição mais correta, deverei retomá-la ao longo do trabalho, principalmente no capítulo 4. A idéia básica é a de que a unidade mínima de formulação manifesta na conversação cotidiana sejam os pequenos agrupamentos prosódicos conhecidos como *unidades entoacionais* (Chafe, 1994), ou *pacotes* (do inglês *parcels*; Clark, 2003).

diversos pontos de completude conforme vai sendo incrementada, mas é a “entoação [que] desempenha um papel preponderante na determinação de quais enunciados sintaticamente completos devem ser projetados pelos receptores como unidades completas” (p. 157).

Indo de encontro com o favorecimento de qualquer uma dessas dimensões de análise, sintática ou prosódica, Selting (2000) argumenta não ser possível estabelecer qualquer tipo de hierarquia entre esses níveis na constituição de TCUs/TRPs, sendo a relação mútua entre eles, sob a restrição de certos contextos interacionais, o fator determinante das unidades na interação. Inicialmente, a autora busca demonstrar que, diferentemente de SSJ e dos trabalhos clássicos da AC, a sintaxe, apenas, não poderia servir como critério de delimitação de TCUs. Para isso, a autora traz dados que revelam, em consonância com Ford e Thompson (1996), que, em muitos contextos, é a contextualização prosódica da sintaxe na fala que permite dizer se um possível ponto de completude sintática foi projetado para ser um TRP de fato (Selting, 2000: 494).

Em seguida, a autora parte para uma segunda análise buscando demonstrar que, em alguns casos, exatamente o oposto ocorreria. Analisando uma série de cinco unidades entoacionais sucessivas marcadas por reparos e reformulações (“wir ham in diesem semester Einige Auftritte gehabt/Auch ähm/Überwiegnmuss ich sagen/Anlässlich ähm/ jaa des streiks der gewesen is”),⁴⁶ a autora afirma sobre a passagem: “A sentença é empacotada em diferentes unidades entoacionais; cada uma inicialmente parece sinalizar o início de uma nova sentença e TCU, mas cada uma acaba sendo a continuação sintática da sentença complexa previamente iniciada” (p. 502). E complementa, afirmando que, embora seja produzida de maneira hesitante, “as projeções sintáticas que começam e continuam a cada grupo prosódico sucessivo são relativamente fortes, de modo que a oração é ouvida como tal, e a unidade inteira como uma única TCU” (p. 502). A evidência empírica que a autora traz é um sinal de acompanhamento (“mhm”) que o então-receptor reservou para o momento em que toda a passagem é concluída.

Essa última análise de Selting, porém, me parece problemática tanto do ponto de vista teórico quanto empírico. Teoricamente, a análise nos leva a um contra-senso em

46 A tradução oferecida pela autora é: “We did give several performances this semester/ also uhm/ mainly I must say/ occasioned by uhm/ well the strike that was going on”.

relação às premissas da AC: se a TCU deve ser definida, como propõe a autora, como sendo “a menor unidade lingüística completa *de relevância interacional* em seu contexto particular” (p. 512), a conclusão dessa análise seria a de que as unidades prosódico-sintáticas mínimas que constituem esse pedaço de fala, não sendo TCUs, não possuiriam relevância interacional intrínseca. Empiricamente, a autora ignora a significação de outras evidências presentes ao longo da passagem, como os “recursos de retenção de turno” (e.g. “ähm”, alongamentos de sons, e pitch mediano ao término de contornos prosódicos), embora ela própria ateste a presença desses recursos (p. 500). Não está nada claro para mim por que motivo um falante empregaria recursos para “retenção do turno” em momentos que não se mostrem relevantes para a tomada do turno pelo então-receptor.⁴⁷

O fato de a passagem como um todo corresponder a um esquema abstrato complexo (grosso modo equivalente a uma sentença) instanciado por meio de várias unidades entoacionais revela, a meu ver, o caráter fundamentalmente local do processo de formulação. Desenvolvendo-se em uma unidade entoacional por vez, esse processo abre espaço para a ruptura, a reorientação e o re-acabamento por parte do produtor – o princípio do “ajuste ao interlocutor”, intrínseco ao funcionamento da conversação, que é discutido por SSJ – e para a reação em potencial desse interlocutor a cada novo incremento produzido, cada um deles em si uma nova contingência da fala. Considerar o sinal de recepção ao final da “sentença completa” como evidência em si de um TRP me parece uma consideração analítica por demais presa a um dado específico e, como argumentado acima, passível de refutação tanto em termos teóricos quanto empíricos.

Como um último ponto a ser destacado, Ford et al. (1996) chamam a atenção para o fato de as análises sobre projetabilidade deverem considerar a produção simultânea de múltiplas trajetórias, não apenas definidas por recursos léxico-gramaticais e prosódia, mas também por gestos e seqüências interacionais mais amplas. Apesar disso, análises que mostrem como cada uma dessas dimensões de uso interagem para

⁴⁷ Jefferson (1984) destaca que os locais nos turnos que apresentam problemas de progressividade (i.e. disfluências de vários tipos) se mostram empiricamente propensos às sobreposições de falas – embora a própria autora também não veja esses locais como pontos de relevância para transição (ver também Schegloff, 2002: 297).

produzir múltiplas coordenadas de ação em contextos interacionais específicos ainda são raras (mas ver Goodwin, 2002; uma análise que me parece caminhar nesse sentido).

3.5. Uma proposta sobre os conceitos-chave⁴⁸

A discussão anterior trouxe algumas das contribuições mais importantes no campo da AC para a reflexão sobre o fenômeno da projetabilidade na conversação – fenômeno crucial para o entendimento da organização da interação social e, reflexivamente, da organização da própria língua. Em SSJ, vimos que o conceito de TCU aparece atrelado ao conceito de TRP, isto é, que a cada possível término de uma unidade de construção de turnos um local de relevância para transição de falantes se sucede. Essa projeção de TRPs, contudo, seja no âmbito de unidades discursivas menores (unidades gramaticais como sentenças) ou maiores (atividades estruturadas como histórias ou listas) não deve ser confundida com “determinação”; a interação é fundamentalmente caracterizada pela contingencialidade das ações dos participantes diante de cada novo incremento de fala, e, em vista disso, projetos estabelecidos num dado momento podem acabar tendo o seu curso alterado, reorientado ou mesmo abandonado completamente.

Os estudos discutidos também sugerem que essa indeterminação na projeção dos possíveis pontos de completude de ações sociais resulte em grande medida da complexa interação entre múltiplos recursos lingüísticos e/ou não lingüísticos na construção da fala. Inicialmente, SSJ deram primazia à sintaxe em suas análises, mas vários pesquisadores que se seguiram a eles apontaram o papel fundamental de outras dimensões da língua, como a gestualidade e principalmente a prosódia. Esses diferentes recursos, quando interagindo, projetam múltiplas trajetórias de ação simultaneamente e há evidências de que a convergência dessas trajetórias seja o melhor fator preditivo dos pontos em que a transição entre falantes de fato opera.

⁴⁸ Tendo em vista que não disponho de um corpus transcrito e analisado de uma língua familiar, como o português, optarei, na reflexão que se segue, por fazer uso de situações hipotéticas de conversa. Embora eu saiba que a teorização sobre exemplos fabricados é contrária à orientação metodológica do campo da AC, acredito que essa estratégia seja justificável no presente contexto, uma vez que meu atual foco de análise empírica é sobre uma língua ainda carente de ser descrita e cuja segmentação em unidades é ainda uma questão incipiente. Assim, a proposta aqui desenvolvida deve ser lida como um *exercício de reflexão*, que se faz necessário para a articulação teórica deste trabalho e que, futuramente, poderá ser devidamente confrontado com dados, não apenas do português, mas, espera-se, da própria língua de sinais brasileira.

Diante dessas reflexões, podemos agora passar a uma formulação explícita desse tema no presente trabalho, em especial no que diz respeito à definição dos conceitos de turno, TCU e TRP. A essência de minha proposta é a de que a dificuldade nas análises sobre o fenômeno da projetabilidade na conversação espontânea se deva, em grande medida, a uma visão um tanto inflexível (ou ainda “clássica”) sobre a natureza dessas categorias. Esse tema exige, diferentemente, uma compreensão dos termos em questão considerando-se diferentes graus de participação na interação (em relação às categorias “turno” e “TCU”) e diferentes graus de relevância para transição de falantes (em relação à categoria “TRP”). Tal proposta irá destacar alguns aspectos da formulação de SSJ que carecem de esclarecimento e/ou reformulações, embora a essência da proposta desses autores mantenha-se a mesma.

O argumento central pode ser assim elaborado: se turno, TCU e TRP são categorias com as quais os participantes de uma conversa tacitamente operam para a realização de ações práticas, é de esperar que essas categorias apresentem o tipo de abertura e flexibilidade necessárias para a sua operacionalidade na interação. Uma proposta teórica sobre categorias que parece atender a esses requisitos, isto é, que se revela não apenas cognitivamente mas também interacionalmente plausível, é a teoria dos protótipos (Rosch, 1978). Embora essa teoria tenha sido devidamente criticada por tratar a categorização de uma maneira descontextualizada e idealizada (Billig, 1987; Potter e Wetherell, 1987; cf. Edwards, 1991), Edwards (1991) argumenta de maneira persuasiva que ela é um importante avanço em relação à teoria clássica, indo parcialmente ao encontro das teorias centradas na interação pelo fato de abrir espaço para a *indicialidade*⁴⁹ do processo de categorização no uso da língua na interação espontânea.

⁴⁹Para Garfinkel (1967), uma das propriedades essenciais da língua é a *indicialidade* (*indexicality*), de acordo com a qual as expressões empregadas na interação cotidiana não podem ter o seu significado definido a menos que aquele que a interprete conheça ou pressuponha algum aspecto particular da situação imediata. Nos estudos lingüísticos, essa propriedade costuma ser atribuída a categorias específicas da língua, como os pronomes, mas Garfinkel argumenta que a propriedade permeia toda e qualquer situação de uso das línguas naturais.

Influenciada por Garfinkel, então, a perspectiva da AC trata do fenômeno de categorização como um trabalho ad-hoc que é feito pelos participantes sempre de maneira situada e com o intuito de atingir determinados propósitos práticos – e não como um processo racional e inocente de se fazer sentido do mundo, independentemente de contextos específicos.

Um exemplo ilustrativo dessa visão indicial do processo de categorização é trazido por Heritage (1984: 149, citado em Edwards, 1991): “[I]f, in a pencil and paper test situation, I am asked to ‘name

A proposta central da teoria dos protótipos, amplamente investigada em pesquisas voltadas para a língua, a percepção e a cognição (Lakoff, 1987), é a de que diferentes elementos são agrupados sob uma categoria, não por compartilhar um conjunto intrínseco de traços necessários e suficientes, mas, diferentemente, por apresentar traços familiares (*family resemblances*) uns em relação aos outros. Isso significa que, numa seqüência de elementos de A a F de uma categoria hipotética, em que A lembra B, que lembra C, que lembra D, que lembra E, que lembra F, é possível que A e F sejam tão dessemelhantes em várias de suas características que pareçam pertencentes a categorias distintas, embora não o sejam. Ao mesmo tempo, certos elementos de uma categoria sempre se destacam como o seu melhor exemplo, ou o mais representativo: os chamados protótipos. A coexistência de elementos prototípicos e elementos que deles se distanciam em diferentes graus faz com que as fronteiras de muitas categorias apresentem limites difusos.

Mostrando a relevância dessa proposta para os estudos centrados na interação, Edwards (1991) destaca que

[t]he idea that semantic categories have fuzzy membership boundaries, inequities of membership and permit multiple and even contrasting possibilities for description suggests that language's category system functions not simply for organizing our understanding of the world, but for *talking* about it in ways that are adaptable to the situated requirements of description, and to differences of perspective, and to the need to put words to work in the pragmatics of social interaction. Further, because categorical descriptions involve choice, and are rhetorically consequential, they also potentially display the speaker as positioned, interested and accountable in a loosely moral sense for how things are described, and for the interactional consequences of descriptions (p. 523, itálicos no original).

Embora Edwards esteja discutindo categorias de natureza semântica mais evidente, como as palavras, há amplas evidências de que essa mesma forma de funcionamento possa ser estendida a categorias mais estruturais, como por exemplo as construções sintáticas (Langacker, 2000: 99). Tendo essa questão teórica como pano de fundo, então, podemos tecer algumas considerações sobre os conceitos de turno, TCU e TRP, recuperando, sempre que relevante, os trabalhos já introduzidos que se voltaram para essa questão.

some typical birds', I may very likely mention robins and sparrows. ... But neither is at all likely to come to mind when I am greeted at the door with: 'I've just put the BIRD in the oven'."

3.5.1. Turnos e TCUs

Embora até o presente momento apenas os conceitos de TCU e TRP tenham sido foco de discussão, uma reflexão adequada sobre o fenômeno da projetabilidade exige também uma discussão sobre o conceito de turno. O motivo é o de que, para um pesquisador como eu, que se vê na necessidade de abordar uma língua cuja organização interna em unidades mínimas é desconhecida, a primeira porta de entrada para análise dos dados são os turnos. A definição de turnos como “a vez” de cada interlocutor na interação, porém, em geral assumida nos trabalhos sem uma reflexão explícita, se mostra por demais vaga e insatisfatória. Muitos pesquisadores, por exemplo, não consideram certas condutas gestuais (i.e. acenos de cabeça ou mesmo gestos icônicos ou metafóricos de conteúdos substantivos claros) como instâncias de turnos, argumentando que o sistema de tomada de turnos de SSJ foi desenvolvido pensando-se na organização de turnos-*na-fala*⁵⁰ – a fala entendida nesse argumento em seu sentido estrito, de comunicação verbal.

Minha visão, porém, é a de que o turno seja uma fala verbal e/ou não-verbal que realiza uma contribuição para a interação em curso na medida em que se mostra seqüencialmente implicativa. A discussão dessa definição será iniciada pela última parte, referente à função dos turnos, reservando a reflexão sobre a sua natureza formal para o momento seguinte.

A implicação seqüencial de um turno está relacionada tanto ao caráter retroativo das participações (i.e. uma participação *demonstrando* orientação ao que a precedeu) quanto ao seu caráter prospectivo (i.e. uma participação *demandando* orientação do que segue a ela). Esse tipo de implicação é, em essência, o que SSJ identificam empiricamente: “Os turnos exibem características organizacionais evidentes que refletem suas ocorrências em uma série. Eles geralmente têm uma estrutura de três partes: uma que aborda a relação de um turno com um anterior; uma envolvida com o que está ocupando o turno, e uma que aborda a relação de um turno com um posterior” (p. 43).⁵¹

⁵⁰ Gene Lerner, em resposta à discussão sobre gestualidade e a organização da troca de turnos, na lista “Languse”, em outubro de 2007. A resposta do autor pode ser acessada em <http://www.list.hum.aau.dk/pipermail/languse/Week-of-Mon-20071022/002831.html>.

⁵¹ Vários exemplos em que esses direcionamentos são incorporados explicitamente no desenho do turno são trazidos por SSJ (p. 43-44; ver também Schegloff, 1996: 61-64), como o que se segue abaixo:

Um rápido olhar sobre os próprios dados trazidos pelos autores revela, porém, que tal descrição (que, poderíamos dizer, reflete uma visão prototípica da estrutura de um turno, com cada parte separadamente designada para a realização de uma função específica) nem sempre se manifesta explicitamente nos turnos-na-fala. Por exemplo, vários turnos tendem a fundir as funções retrospectivas e prospectivas em uma única unidade, ou, apesar da pressão do sistema de tomada de turnos, a fragmentá-las ao longo de várias unidades. Nesses casos menos prototípicos, os turnos são construídos de tal maneira que algumas dessas funções podem até ser ofuscadas frente aos olhos do analista. É de esperar, portanto – seguindo o argumento de Edwards sobre o trabalho de categorização –, que as diferentes opções de estruturação de turnos em uma ou mais “partes” estejam relacionadas a determinados *trabalhos interacionais*⁵² que os participantes realizam em face dos requerimentos da situação por meio da escolha de uma categoria em meio a um repertório de opções.

Há evidências de que isso de fato acontece. Por exemplo, bastante distantes do turno prototípico tripartido estão os turnos formados por “continuadores”, do tipo “mh”, “mh-hm”, “ãhn”, etc (Schegloff, 1982). Quando esse tipo de turno é empregado, a participação de um interlocutor pode ser mais bem interpretada como envolvendo um *papel primário receptivo*, de acordo com o qual a palavra é ratificada por ambos os interlocutores como pertencendo ao falante-corrente – como tende a acontecer, por exemplo, durante a narração de uma história. Nesses contextos, o trabalho interacional do receptor que produz um continuador (i.e. de demonstrar que a palavra permanece com seu interlocutor, a despeito de sua contribuição) é realizado, em parte, pela escolha de um tipo de turno que salienta a sua função retrospectiva e ofusca a sua função prospectiva.⁵³ Nesse sentido, os continuadores podem ser interpretados como

A: I mean you went- you went through a- a long rap on that one

→ B: [Yeah_i],[so I say that would bum you out then_{ii}],[hunh_{iii}].

Nesse exemplo, cada uma das partes do turno inserida entre colchetes revela, respectivamente, i) a sua orientação retrospectiva, ii) a contribuição do turno per se, e iii) a orientação prospectiva; cada uma delas incorporada explícita e separadamente no desenho do turno.

⁵² A noção de “trabalho interacional” está intimamente ligada ao conceito de “formulação”. A idéia é a de que toda fala realiza, ou contribui para a realização de uma determinada ação social, e que o resultado final dessas ações, em geral disponível ao analista a posteriori, deva sempre ser entendido como uma conquista (*achievement*) dos participantes.

⁵³ Previamente ao trabalho de Schegloff (1982), por exemplo, a função prospectiva dos turnos do tipo “mh” e “mh-hum” tinha sido amplamente ignorada pelos pesquisadores, que atribuíam a esses usos funções tais como a de “demonstrar entendimento sobre a fala do interlocutor”. Schegloff revelou, nesse

pertencentes a uma subcategoria (da categoria mais ampla dos turnos) desenhada especificamente para a realização de ações sociais específicas; ela mesma podendo apresentar elementos mais ou menos prototípicos relacionados a funções também diferenciadas.

Por outro lado, considerando-se a ausência de uma relação biunívoca entre forma e função na interação, é o olhar para a fala situada que pode mais bem revelar o trabalho que diferentes instâncias de turnos irão desempenhar num dado contexto imediato. Como aponta Edwards, a tipicidade de uma categoria não é dada aprioristicamente e aplicada unilateralmente, mas está submetida a um contínuo processo de interpretação e negociação por parte dos participantes (cf. exemplo sobre “a *ave* no fogão”, na nota de rodapé 49). Um procedimento de verificação da análise sobre a implicação seqüencial dos turnos, então, provém da própria natureza dialógica da fala-em-interação. Como propõem SSJ, analisando-se a seqüência triádica X-Y-Z, entende-se que as falas X e Z do(s) próprio(s) co-participante(s) ajudem a revelar o tipo de trabalho ao qual uma dada fala Y se orientou e o trabalho que ela própria realizou no contexto imediato da interação (p. 52-53).⁵⁴ Desse modo, independentemente da possível associação de certos elementos com certos tipos de trabalho que eles desempenham (e.g. o “mh-hm” e a concessão de um turno mais extenso ao falante-corrente, que continua como mantenedor da palavra), deve-se ter em mente que tais correlações são, em última instância, dependentes do processo imediato de interação entre os participantes.⁵⁵

sentido, a orientação prospectiva implícita nesse tipo de turno por meio da análise de como os participantes reagem diante dele. Ele mostrou que o falante-corrente, em face dessa contribuição, segue falando em parte por saber que um turno mais extenso continua sendo ratificado por seu interlocutor. A evidência dessa orientação pode ser mais bem observada quando tais expectativas são quebradas, como por exemplo quando, ao término de uma piada, o interlocutor reage com um “mh” ao invés da risada esperada, ao qual o piadista poderia plausivelmente retrucar, diante do fracasso de seu empreendimento: “É isso! A piada acabou!”.

⁵⁴ SSJ referem-se a esse procedimento de análise como um “procedimento de prova” (*proof procedure*), o que parece superestimar o seu valor heurístico na análise da conversa (Coulter, 1983).

⁵⁵ Refletindo sobre o mesmo caso hipotético da piada discutido na nota de rodapé 53, podemos pensar em como a produção do continuador ao término da piada poderia estar submetida a um processo contingente de negociação sobre o seu estatuto categorial, tanto por parte do receptor quanto do produtor.

O receptor, por exemplo, ao produzir um continuador ao final da piada, poderia estar de fato, inocentemente, esperando algo mais por vir, mostrando-se incapaz de identificar o clímax já apresentado; ou poderia, uma vez tendo identificado um clímax totalmente aquém de suas expectativas, empregar o continuador ironicamente, sinalizando ao narrador, pelo contraste entre a sua participação de fato e o tipo de participação esperada, a falta de graça da piada. Similarmente, o produtor poderia interpretar o

Tal abertura e flexibilidade, cruciais para a operacionalidade da conversação cotidiana, só são possíveis porque categorias como “turno”, ou “turno-continuador”, por exemplo, apresentam efeitos de tipicidade, com algumas instâncias se revelando mais prototípicas e outras menos prototípicas num contínuo gradativo dependendo do uso que delas é feito em situações específicas. Essa flexibilidade não seria possível dentro de uma visão rígida e discreta de categorias (a visão clássica), segundo a qual os elementos seriam categorizados de acordo com traços necessários e suficientes, definidos aprioristicamente com base nas características supostamente inerentes a esses elementos.

Na literatura, dentre as abordagens que a meu ver se assemelham a essa proposta, está o trabalho de Ford e Thompson (1996), que consideram como instâncias de trocas de falantes não apenas o que elas chamam de turnos plenos (*full turns*), mas também os turnos de retorno (*backchannel turns*).⁵⁶ Ao passo que o primeiro tipo envolve o controle da palavra, o último tipo envolve os já referidos enunciados em que o participante assume um papel primário de receptor. A própria nomenclatura adotada pelas autoras, fazendo referência a turnos *plenos*, sugere a idéia de que certos tipos de turnos realizem uma contribuição interacional de maneira mais prototípica (i.e. participações que envolvem não apenas a tomada do *turno*, mas a tomada da *palavra*, apresentando funções retrospectiva e prospectiva explícitas), considerando-se o papel operacional que toda contribuição deve, de alguma maneira, desempenhar na interação.⁵⁷

continuador de seu interlocutor inocentemente, afirmando como na nota anterior “É isso! A piada acabou!”, ou poderia reagir à mesma situação interpretando o turno de seu interlocutor como uma provocação com função prospectiva explícita, podendo retrucar algo como: “O problema é seu, que nunca entende uma piada!”.

A opção por uma ou outra interpretação sobre o papel do turno irá depender, portanto, do olhar para a troca de falas imediata entre os participantes, que irá dizer como eles próprios categorizaram o turno em questão, isto é, como *formularam* a interação imediata (Sacks, 1992).

⁵⁶ O termo *backchannel* foi criticado por conferir um papel excessivamente passivo ao interlocutor e ignorar a função prospectiva que caracteriza os turnos formados por continuadores (Schegloff, 1982). Apesar disso, as autoras parecem empregar o termo do mesmo modo como ele é tratado por Schegloff, e não em seu sentido tradicional.

⁵⁷ Entendo, seguindo SSJ, que toda contribuição interacional deva, de algum modo, apresentar orientações retrospectivas e prospectivas tendo em vista que tanto o estabelecimento quanto a manutenção das interações – ou *encontros*, nas palavras de Goffman (1961) – envolvem a necessidade contínua de *ratificação mútua de um foco cognitivo e social*. Isso não impede que certas contribuições rompam com esse padrão, mas explica porque falas que violam esse princípio tendam a ser mal interpretadas em seu contexto e/ou a se tornarem questionáveis pelos participantes (ver Sacks e Schegloff, 1973, que mostram

SSJ, por sua vez, apresentam uma caracterização empírica dos turnos que deixa clara a sua função, mas não abordam a questão dos diferentes graus de participação ou envolvimento do interlocutor dependendo dos tipos de turnos que ele opta por empregar (e.g. uma participação plena, que envolva a tomada do turno *e também* da palavra, explicitando o caráter retrospectivo e prospectivo da participação; ou uma participação mais comedida, que não envolva a tomada da palavra e obscureça o seu caráter prospectivo). Desse modo, à caracterização original de SSJ eu acrescentaria a idéia de que, embora sejam “regularmente estruturados” em três “partes”, os turnos são de maneira geral “imbuídos” de três “funções”, que podem ser manifestas estruturalmente de diferentes formas. Num extremo, como em turnos que compõem histórias, essas funções podem estar distribuídas ao longo de várias unidades; em outro extremo, como no caso dos continuadores, elas podem aparecer fundidas numa única unidade minimalista; e em casos mais prototípicos, elas podem se desdobrar em três partes separadas e claramente identificáveis. A opção por esses diferentes níveis de explicitação das funções dos turnos deve estar, como argumentado por SSJ, submetida às restrições do sistema de tomada de turnos, e, como argumentado aqui, relacionada a certos tipos de trabalho interacional levado a cabo pelos participantes em face de situações específicas.

Até aqui, a discussão sobre o conceito de turnos envolveu aquilo que, poderíamos dizer, constitui o “pólo semântico” do termo de acordo com a definição proposta (i.e. uma contribuição para a interação que se mostra sequencialmente implicativa). Uma outra dimensão da definição apresentada que precisa ser discutida, então, refere-se à sua natureza formal, isto é, ao modo adequado de caracterizarmos o seu “pólo fonológico”. Tendo em vista que os turnos são constituídos por uma ou mais unidades de construção, o que entra em jogo aqui são considerações a respeito dos recursos formais que são utilizados para a construção das chamadas TCUs.

De acordo com a definição aqui proposta, TCUs (e, conseqüentemente, turnos) são formados por recursos verbais e/ou não verbais, isto é, podem envolver a exploração de recursos fonéticos, lexicais, prosódicos, sintáticos e também gestuais. Porém, apenas a constatação de que recursos de diversas naturezas são empregados na construção de

um tipo de trabalho interacional específico que precisa ser realizado nos encerramentos de conversas em vista da necessidade de se romper com essa implicação sequencial).

TCUs não resolve o problema de sua definição. Por exemplo, tendo em vista o fato de que a postura responsiva (ativa) de um interlocutor pode envolver desde uma mudança sutil de postura, expressão facial ou olhar, até o início de uma fala altamente competitiva, acompanhada de gesticulação, *pitch* e volume acentuados, que nível de envolvimento lingüístico e corporal dentro desse contínuo será necessário para que consideremos uma dada conduta uma (tentativa de) tomada de turno? Uma mera mudança de olhar e de expressão facial durante um turno-em-curso pode, em determinadas situações, revelar um grau menor ou maior de implicação seqüencial, suscitando uma reação do falante em termos de re-calibrações em sua fala ou mesmo questionamentos explícitos. Seriam esses casos instâncias de turnos e estariam eles submetidos ao mesmo princípio do um-de-cada-vez elaborado por SSJ?

A proposta de que o discurso espontâneo estrutura-se por meio de entidades autônomas, isto é, de *unidades* de construção de turnos, foi inicialmente elaborada por SSJ. Para os autores, a orientação dos participantes a essas unidades e aos seus possíveis pontos de completude, distribuídos *de maneira discreta* ao longo da fala, seria o principal fator responsável pela coordenação da troca de turnos na fala-em-interação. Essas unidades seriam formadas principalmente por recursos lexicais e sintáticos, embora a prosódia tenha sido também apontada como capaz de redefinir a estruturação desses pontos de completude em contextos específicos.

Cabe aqui lembrar que tal proposta emergiu num contexto bastante enviesado em favor da identificação de unidades com essas características. Em primeiro lugar, os dados utilizados envolviam principalmente conversas telefônicas, uma modalidade de comunicação que, por eliminar o canal visual natural da interação face-a-face, pressiona os participantes a se apoiar com mais ênfase sobre os recursos de vocalização para o gerenciamento da interação (para uma análise sobre uma situação similar, ver Clark, 2005). Tal vocalização, diferentemente das articulações gestuais que subjazem a ela ou que co-ocorrem com ela, é um elemento que pode ser satisfatoriamente delimitado em seu início e fim pela identificação de pausas, isto é, momentos de ausência de vocalização. Em segundo lugar, parece-me possível identificar um viés da escrita nos trabalhos da tradição de SSJ, tendo em vista que a análise não se debruça sobre a gravação diretamente, sendo sempre mediada pela transcrição escrita dessa gravação, com sua distribuição espacial de unidades mais comportada, independentemente das convenções buscadas para aproximar a transcrição da fala espontânea. Uma das

consequências mais diretas desse viés foi a subestimação do importante papel que a prosódia desempenha no discurso espontâneo, e uma significativa escassez de discussão sobre o possível papel que os gestos teriam na construção dessas unidades.⁵⁸

Por um lado, é preciso reconhecer que o favorecimento daqueles recursos lingüísticos codificados na escrita é um primeiro passo natural na teorização sobre a ordenação da conversa espontânea, tendo em vista a importância indiscutível da tecnologia da escrita como ferramenta analítica e o momento histórico dentro do qual a AC emergiu. Por outro lado, permanecer apenas nesse âmbito significa deixar de lado uma grande riqueza de recursos que os participantes de fato utilizam na interação face-a-face – um meio que, no próprio campo da AC, é tido como primordial para o estabelecimento da ordem social na espécie humana (Schegloff, 1996). Nesse sentido, em especial diante das tecnologias hoje disponíveis, a subestimação maior ou menor da prosódia e da gestualidade do escopo de análise da AC sobre a sistemática da organização da tomada de turnos me parece uma atitude tão arbitrária quanto tem sido, por parte dos lingüistas, a exclusão desses recursos do escopo de análise sobre o sistema lingüístico – como argumentam Langacker (2001: 147), McNeill (1992), entre outros.

Com relação à gestualidade, uma reflexão adequada sobre o seu papel na construção dos turnos exige a devida discriminação dos diferentes tipos de gestos existentes, com suas características formais e funcionais peculiares.⁵⁹ Por exemplo, as *batidas* estão intimamente associadas ao ritmo da fala, e observações naturalísticas me fazem acreditar que elas possam estar correlacionadas com unidades entoacionais,

⁵⁸ Embora seja certamente incorreto dizer que as atividades não-verbais não tenham sido levadas em consideração no campo da AC, parece-me também inegável que os aspectos vocais da produção foram freqüentemente tratados sem qualquer referência à prosódia e aos gestos. Embora essa restrição possa parecer perfeitamente aceitável para pesquisadores que cotidianamente se defrontam com dados de LOs, quando passamos a considerar a interação em LSs nos vemos obrigados a questionar em que medida essa cisão é adequada ou mesmo relevante.

⁵⁹ McNeill (1992), considerando vários trabalhos sobre gestualidade inclusive o seu próprio, propõe a seguinte classificação: *gestos dêiticos*, como apontamentos, indicam referentes no espaço físico imediato ou em locais específicos no espaço onde certos referentes são alocados no decorrer do discurso; *gestos icônicos* representam idéias concretas, como por exemplo as mãos abertas, ligeiramente curvadas e direcionadas umas às outras podem se referir à “bola”; *gestos metafóricos* representam idéias abstratas, como num dos exemplos trazidos pelo autor, em que um gesto similar à ação de “segurar um pacote” é realizado por um narrador ao se referir ao “cartum” que será contado; *batidas* são gestos manuais, em geral formados por movimentos para baixo e para cima, que acompanham e ajudam a compor o ritmo da fala; e os *emblemas* são gestos altamente convencionalizados numa dada comunidade, tanto em sua forma quanto em seu significado, tal como o “jóia”, o “dar-de-ombros” e o “mais-ou-menos”, que fazem parte de nosso repertório lingüístico cotidiano.

apresentando diferenças qualitativas de configuração e movimento de uma unidade para outra.⁶⁰ Nesse sentido, seria possível que a observação das batidas fosse um recurso adicional para identificação de unidades entoacionais.⁶¹ Outros tipos de gestos podem integrar o processo de construção de um turno adicionando-lhe camadas de significação, ocorrendo simultaneamente à “fala” ou ainda em locais onde, a princípio, esperaríamos a fala estrito senso. É o caso dos dêiticos, emblemas, gestos icônicos e metafóricos. Há evidências sólidas para se acreditar que esses diferentes tipos de gesto estejam relacionados com a gramática (McNeill, 1992) e, conseqüentemente, com a coordenação da ação na conversação (Goodwin, 2002: 39-41), não havendo razão para considerá-los, portanto, elementos de fora do escopo do estudo da fala-em-interação.

O que à primeira vista parece mais complicado é considerar em que medida o olhar, as expressões faciais, e as mudanças de orientação e postura da cabeça e do corpo poderiam integrar as TCUs e o processo de construção de turnos. O problema central, nesse caso, refere-se à delimitação dos turnos, tendo em vista o princípio do um-de-cada-vez. Como argumentado acima, tal princípio foi formulado tomando-se por base a idéia de troca de turnos-*na-fala* – a fala se apresentando nos dados de análise em sua forma mais restrita, relacionada aos recursos de vocalização que aparecem numa transcrição. Nesse sentido, estender o conceito de fala para abarcar as diferentes formas de gestualidade não comprometeria esse princípio e, conseqüentemente, o mecanismo proposto para a organização da troca de turnos?

Por exemplo, é comum algumas pessoas acenarem continuamente com a cabeça enquanto o interlocutor fala. Seria essa uma instância de sobreposição de falas, sendo uma dessas “falas” de natureza estritamente gestual? Certamente o aceno de cabeça é uma contribuição à interação, com implicações seqüenciais possivelmente similares às dos continuadores em contextos de sobreposição. Até aqui, não há problemas. É um fato atestado no campo da AC que nem todos os tipos de sobreposições sejam competitivos

⁶⁰ Até onde eu saiba, essa é uma hipótese que ainda não foi investigada. Loehr (2004) aborda a relação entre gesto e entoação com um foco distinto, buscando correlações entre os acentos e o ritmo das unidades prosódicas, de um lado, e os golpes de diferentes tipos de gestos incluindo as batidas, de outro.

⁶¹ McNeill (1992) argumenta que as batidas estão relacionadas à estrutura do discurso, sinalizando mudanças em diferentes níveis da narrativa: a narrativa per se, a metanarrativa e a para-narrativa. Essa análise mostra que esse tipo de gesto se mostra possivelmente relevante, também, para a organização de seqüências maiores de fala.

e (conseqüentemente) problemáticos para os participantes, um deles sendo a sobreposição entre uma fala-em-curso e os continuadores (ver discussão da seção 6.2).

Contudo, parece-me que há vários tipos de olhares, posturas e expressões na fala-em-interação que envolvem graus importantes de participação ou envolvimento, mostrando-se problemáticos para os participantes e repercutindo tanto na fala com a qual co-ocorrem quanto na fala subsequente. Vem-me à memória situações em sala de aula em que meu professor reagia sem palavras às formulações de seus alunos, mas com gestos, expressões e olhares particulares que muitas vezes repercutiam diretamente na fala desses alunos, então acompanhada por disfluências e submetida a reformulações que mostravam clara orientação a esses “turnos gestuais”. Embora alguém pudesse argumentar que o modo de funcionamento do mecanismo de tomada de turnos permaneceu inalterado, na medida em que os próprios efeitos na fala dos alunos revelavam o caráter problemático dessas sobreposições de fala, a mera equivalência dessa situação com a de sobreposição de falas tradicionalmente estudada corre o risco de obscurecer diferenças que podem se mostrar significativas.⁶²

Os trabalhos de Kendon (1972), McNeill (1992) e colegas nos oferecem também subsídios teóricos importantes para analisar a delimitação de TCUs e turnos envolvendo formas de gestualidade tais como as acima referidas. Seus estudos sobre a estruturação interna de unidades gestuais em seqüências e fases – em especial a fase de preparação, a

⁶² Por exemplo, neste contexto de sobreposição de um turno “gestual” com um turno “falado”, minhas observações casuais sugerem que: a) a extensão da sobreposição parecia ser potencialmente mais longa do que as sobreposições de fala comuns; b) o tratamento da sobreposição como algo problemático parecia se manifestar no turno de apenas uma das partes, a saber, a dos alunos que mantinham a fala-em-curso, uma vez que o gesto do professor permanecia basicamente inalterado por períodos relativamente prolongados; e c) o tipo de repercussão na fala desses alunos nem sempre parecia envolver os recursos de resolução de sobreposições descritos na literatura (ver discussão da seção 6.2 sobre o tamanho usual de duração das sobreposições, os seus efeitos sobre a fala dos participantes e os recursos de vocalização empregados para a resolução desse problema interacional). Embora essas observações necessitem de investigações cuidadosas, elas favorecem o apelo de Gene Lerner de que a sistemática da organização da tomada de turnos na interação refere-se, a princípio, a turnos-*na-fala* (estricto senso) e que devemos tomar precauções na ampliação do escopo desses conceitos sem a devida consideração das acomodações que podem se mostrar necessárias para o sistema.

Alguns pesquisadores têm optado por modificações terminológicas (e.g. fala-e-conduta-corporal-na-interação, em Mondada, 2006) a fim de abarcar essas novas dimensões da organização da interação. Minha preferência, contudo – salvo as considerações tecidas sobre o cuidado necessário para a integração de novos recursos aos já existentes –, é a de manter a expressão “fala-em-interação” considerando, em um sentido amplo do termo, tanto a *fala* (envolvendo não apenas léxico e sintaxe, mas também prosódia e gestualidade) quanto a *interação* (e.g. envolvendo não apenas os indivíduos interagentes, mas também recursos físicos e semióticos que os circundam, cf. Goodwin, 2002; Clark, 2005).

fase expressiva, e a fase de retração –, cada uma delas com suas funções particulares, permite segmentar praticamente todas as formas de gestualidade, não apenas as manuais, mas também as que envolvem a cabeça, o olhar e o restante do corpo. Os limites iniciais e finais dessas diferentes fases podem revelar interessantes correlações com a ação dos interlocutores ao longo da conversação, embora essa tenha sido uma questão, até onde eu saiba, ainda inexplorada.

Com relação à prosódia, os estudos no campo da lingüística interacional têm oferecido amplas evidências de seu papel crucial para a construção das unidades da fala-em-interação e para o gerenciamento da troca de turnos (e.g. Couper-Kouhlen e Selting, 1996). Além disso, a observação de que o discurso espontâneo se desenvolve por meio de pequenos agrupamentos prosódicos tem levado vários pesquisadores a conferir à prosódia um papel central no uso da língua, em especial no que diz respeito ao gerenciamento da atenção entre os interlocutores na interação (e.g. Chafe, 1994; Langacker, 2001; Tomasello, 2000a) – tese que será analisada mais detidamente no capítulo 4.

Ono e Thompson (1995), em sua investigação sobre a sintaxe do inglês na conversa espontânea, ajudam a corroborar esse importante papel da prosódia. As autoras apresentam amplas evidências de que, embora os participantes se orientem em relação a esquemas abstratos bastante complexos, a realização progressiva de tais esquemas na conversação se dá num nível bastante localizado (i.e. em termos de sintagmas ou orações simples, encapsulados por unidades entoacionais). Desse modo, embora possamos falar que, em um dado nível atencional, os interlocutores realmente demonstrem orientação a unidades complexas envolvendo cadeias de fala maiores, em última instância os turnos são construídos por meio de um planejamento bastante localizado, uma unidade entoacional a cada vez (p. 235). Os interstícios dessas unidades oferecem pontos de negociação particularmente relevantes aos participantes para a reconstrução de sentidos, seja através de ajustes feitos pelo produtor, seja de intervenções feitas pelo receptor, ambas podendo mudar provisoriamente, ou até mesmo definitivamente, o curso das unidades complexas inicialmente projetadas na fala.

Essa necessidade de orientação local a cada unidade mínima sucessiva, a fim de que complexos esquemas sintáticos e/ou seqüências sejam realizados, mostra que a realização de projetos interacionais mais amplos deva estar subordinada ao processo de produção e interpretação local, realizado por meio da exploração de recursos

prosódicos. Sob essa perspectiva, cada agrupamento prosódico encapsularia um incremento de fala com um estatuto de unidade relevante cognitivamente (i.e. uma unidade de formulação mínima ajustada às restrições de planejamento em tempo real na interação espontânea) e interacionalmente (i.e. uma unidade de formulação mínima ajustada ao interlocutor em face das atuais contingências e dos propósitos imediatos).⁶³ Se essa hipótese de fato estiver correta, poderíamos assumir que as unidades entoacionais constituiriam um dos recursos privilegiados frente aos quais os participantes de uma conversa se orientariam (i.e. direcionariam a sua atenção) a fim de gerenciar e coordenar suas ações sociais na interação.

Retomando a questão original dessa reflexão, então, quanto maior a convergência de recursos formais na produção de uma dada fala (e.g. alta elaboração léxico-gramatical, acompanhada de prosódia e gestualidade marcadas), mais claro será o

⁶³ No uso espontâneo da língua, estudos têm revelado de maneira robusta o caráter pouco complexo das construções gramaticais, com a oração simples constituindo-se a unidade básica de estruturação da fala e apresentando uma maior correlação com as unidades prosódicas (e.g. Croft, 1995, 2005). De maneira geral, atribui-se essa menor complexidade das construções lingüísticas no uso espontâneo da língua (em contraste, por exemplo, com a língua escrita) às limitações de processamento cognitivo do ser humano em face do planejamento da fala em tempo real.

Essa proposta é fortalecida por teses como a da *cognição distribuída*, de Hutchins (2005). Seus estudos etnográficos sobre o processo de uso de ferramentas culturais mostram que o grande desenvolvimento cultural e tecnológico humano emergiu, pelo menos em parte, como solução para nossas limitações de processamento de informação. Segundo o autor, a exploração estratégica de representações estáveis (estruturas materiais naturais ou mesmo culturais com diferentes camadas de significação) viabilizam a operação de cálculos complexos que, de modo natural, poderiam se tornar irrealizáveis ou serem realizados apenas com grande dificuldade. No que concerne ao uso da língua, poderíamos argumentar, por exemplo, que a estabilidade da estrutura possibilitada pelo canal de produção escrita (i.e. a delimitação clara de unidades estruturais por meio de letras, espaços em branco e pontuação) ofereceria precisamente essa estabilidade de representação necessária para a produção e recepção de construções lingüísticas mais complexas do que aquelas encontradas na oralidade.

Independentemente da plausibilidade dessa explicação, porém, parece-me possível vislumbrar também uma motivação social/interacional para esse caráter local e restrito da produção da fala na interação espontânea. O uso da língua oral ou escrita é uma *ação conjunta* de construção de sentidos, realizada a fim de solucionar problemas sociais que interessam a todos os interagentes e que, na condição de *problemas sociais*, não poderiam ser resolvidos unilateralmente, sem a negociação de tópicos e perspectivas. A presença física ou virtual do interlocutor, nesse sentido, exerce uma pressão sobre o processo de construção de sentidos na medida em que as reações sobre as sucessivas formulações se manifestam localmente, momento-a-momento, em tempo real. A interação espontânea face-a-face representa o extremo desse dialogismo, uma vez que o produtor tem a tarefa dupla de planejar e executar a sua formulação ao mesmo tempo em que monitora as reações de seu interlocutor. Nesse sentido, a produção das formulações por meio de pequenos incrementos, negociáveis momento-a-momento em especial pela abertura contínua da fala aos mecanismos de reparo, ajudaria a minimizar as possibilidades de grandes rupturas na “confiança” (i.e. a ordem social) que, como mostra Garfinkel (1963), os participantes se esforçam continuamente para manter.

estatuto pleno de participação do interlocutor na interação; ao mesmo tempo, do outro lado do contínuo, quanto mais minimalista uma dada fala (e.g. formadas exclusivamente por recursos léxico-gramaticais restritos, ou certas formas de gestualidade) mais distante ela estará de uma unidade prototípica de participação na interação. Desde que se mostrem sequencialmente implicativas na conversa, porém, não há razão para excluir quaisquer dessas formas de participação do estudo da organização da conversação.

No que se refere à relevância desses recursos formais para o fenômeno da projetabilidade, vários tipos de gestualidade (e.g. gestos icônicos ou metafóricos que substituem expressões lexicais) parecem poder ser tratados analiticamente sem quaisquer modificações sobre o sistema de tomada de turnos; outros tipos (e.g. expressões faciais, movimentos e posturas corporais) podem eventualmente vir a exigir reformulações, mas não devem se mostrar incompatíveis com o sistema, se é que ele é suficientemente robusto – como, de fato, parece ser.⁶⁴ Com relação à prosódia, tal recurso parece possuir um estatuto privilegiado em relação à sintaxe para o trabalho de coordenação das ações sociais, tendo em vista o processo de produção e recepção mais localizado que caracteriza a conversação espontânea.

Ademais, o caso não parece ser, como propõem SSJ, o de que as “unidades da interação” (tidas como contingenciais) sejam por natureza distintas das “unidades lingüísticas” (tidas como pré-determinadas por constituintes sintáticos estáticos), o que justificaria uma nova terminologia (i.e. TCUs) para a análise da língua em situações espontâneas de uso. Mais apropriado seria, primeiramente, reconhecer que as unidades lingüísticas não estão restritas aos constituintes gramaticais canônicos, podendo envolver também a prosódia, a gestualidade, e inclusive cadeias maiores de fala (Langacker, 2001). Uma vez reconsiderado esse aspecto, a dinamicidade e flexibilidade das unidades que estruturam a conversação podem ser entendidas pelo fato de os turnos-na-fala apresentarem diferentes níveis de construção (i.e. léxico-gramaticais, prosódicos e gestuais), que se relacionam de formas complexas em cada situação imediata e que estabelecem múltiplas coordenadas de participação simultaneamente – daí o seu caráter

⁶⁴ No caso do estudo de uma língua de sinais como a libras, a recusa em considerar a gestualidade como um recurso para a construção de turnos e TCUs simplesmente inviabilizaria qualquer projeto de análise de conversas entre surdos. Ainda que a distinção entre gestos não-lingüísticos (*ad-hoc*) e gestos lingüísticos (convencionalizados) possa ser sustentada para essas línguas (Liddell, 2003a), o fato de ambos serem produzidos pelo mesmo canal de comunicação e de ambos se distinguirem num contínuo (Wilcox, 2004: 134-140) tornaria arbitrária qualquer separação categórica desses conceitos.

inerentemente sensível ao contexto. De maneira geral, porém, o modo estruturado como esses diferentes recursos interagem é um tema que ainda carece de investigação – e a análise de atividades conversacionais como listas e contrastes no capítulo 7 deve caminhar nesse sentido.

3.5.2. TRPs

Tendo em vista a relação íntima entre os conceitos de TCU (ou, como agora parece mais apropriado referir-se, as “unidades mínimas do discurso”) e TRP, a reflexão acima nos conduz a um questionamento similar sobre a noção de “ponto de relevância para transição”. Se, de fato, as práticas de construção de turnos envolvem múltiplas trajetórias de ação que não necessariamente convergem para um mesmo ponto de completude projetado; e se, de fato, quando tais práticas convergem, observa-se empiricamente uma maior relevância para a transição entre falantes do que nos pontos projetados por qualquer nível de análise isoladamente; então seria plausível postular que, assim como se observou na reflexão sobre os turnos e as suas unidades mínimas, os TRPs também sejam passíveis de uma caracterização em termos de pontos com maior ou menor grau de relevância para transição entre falantes, isto é, TRPs mais ou menos prototípicos.

Neste ponto, é interessante recuperar a reflexão sobre os contextos de atividades sociais mais extensas, como a narração de histórias. Pesquisadores como Houtkoop e Mazeland (1986) argumentam que, uma vez ratificada uma narrativa, os TRPs a princípio vinculados às unidades mínimas internas à narrativa seriam bloqueados até que uma unidade identificável como final fosse produzida. O fato, contudo, é que a realização de trabalhos interacionais como os prefácios das histórias não garante a realização do projeto: TRPs não são suspensos ou bloqueados; *apenas têm a sua força diminuída*. Ao ratificar um projeto de fala maior, o então-receptor se coloca numa posição moralmente contestável para interromper o andamento do projeto que ele próprio legitimou. Embora tal fato contribua para que ele *restringa* a sua participação durante a atividade a práticas que não coloquem essa posição moral em dúvida, isso certamente não *impede* uma participação mais plena, que envolva a disputa da palavra.⁶⁵

⁶⁵ Garfinkel e Sacks (1970: 344-345) argumentam que a natureza indicial da língua faz com que qualquer “fato” cuja explicação se torne relevante numa dada interação não possa ser esgotado em tantas quantas forem as palavras empregáveis para explicá-lo. Ao “explicar” um “fato”, cada novo incremento formulado – uma prática de glosagem (*glossing practice*), na palavra dos autores – torna-se

A evidência dessa indeterminação está nas análises que, focando menos a estrutura do que a contingência, mostram as enormes dificuldades por que passam alguns interlocutores para a implementação de projetos de fala maiores (e.g. Goodwin, 1986).

Em outras palavras, é o fato de os TRPs se sucederem a cada unidade mínima, independentemente de contextos particulares, que permite ao mecanismo proposto por SSJ dar conta dos dados empíricos, e, teoricamente falando, que confere à estrutura organizacional da conversa espontânea a flexibilidade e abertura de que ela necessita para a realização de ações práticas em face da imprevisibilidade das contingências interacionais. É essa organização, além disso, que força os participantes, quando desejosos de levar a cabo projetos de fala maiores, a realizar trabalhos interacionais específicos (e.g. o prefácio de histórias e piadas discutido por Sacks, 1974) na esperança de que seus projetos possam ser realizados em sua plenitude.

As histórias, por envolver seqüências de fala mais extensas, exacerbam os efeitos das múltiplas (e muitas vezes contrastantes) projeções de TRPs, cada um deles com diferentes graus de relevância. Contudo, é possível identificar múltiplas trajetórias de ação num nível bem mais restrito de seqüências conversacionais, como por exemplo em seqüências convencionalizadas na forma de esquemas gramaticais complexos. Considere, por exemplo, o diálogo hipotético abaixo, envolvendo a construção condicional “se + então”, altamente convencionalizada no português:

B: E aí? Como você vai fazer amanhã?

→ A: Se eu for na minha mãe amanhã, [eu-

→ B: [Você vai na sua mãe amanhã?! Mas a gente já tinha combinado de ir na casa do Pedro almoçar.

A: Sei, mas acho que dá tempo. Passo lá cedinho e depois vou pro Pedro.

É que faz tempo que não vejo meus pais.

inescapavelmente um novo “fato” da interação, isto é, uma nova contingência da interação, em si passível de explicação. O trabalho infinito da prática da glosagem é restringido apenas na medida em que se satisfaçam os propósitos práticos dos participantes.

Assim, é crucial que o mecanismo de organização da conversa cotidiana ofereça possibilidades estruturais de participação em face de cada uma dessas contingências, de cada um desses incrementos. Mesmo quando inseridos em práticas de glosagem mais amplas (e.g. histórias), a contribuição de cada novo incremento de fala não será a de esclarecer e simplificar o fato até que ele se esgote, mas a de elaborar cada vez mais as circunstâncias que ela própria procura esgotar, o que torna esses incrementos potencialmente abertos ao questionamento pelo interlocutor.

Nesse contexto hipotético, o falante A projeta, na linha assinalada pela flecha, uma seqüência de fala mais extensa. A fala projetada pode ser entendida como “extensa” na medida em que ultrapassa o escopo de uma oração simples expressa numa única unidade entoacional.⁶⁶ A projeção, realizada por meio de uma combinação de recursos léxico-sintáticos (“Se” + tempo subjuntivo) e prosódicos (pitch terminal levemente ascendente, representado pela vírgula), abre espaço para uma ou mais unidades que deverão, ou incrementar a formulação da prótase, ou passar diretamente à formulação da apódose, que por sua vez também poderá ser antecedida e/ou realizada por uma ou mais unidades. Essa projeção, contudo, acaba não sendo concretizada, diante da contingência do questionamento de B assinalada pela flecha – tal questionamento constituindo-se, ele próprio, numa resposta à contingência que a prótase formulada por A introduz na interação. Em outras palavras, embora um TRP forte, projetado pela convergência das construções léxico-gramaticais e prosódicas, ainda não tivesse sido alcançado, o final do contorno prosódico após “amanhã” em si encerrou a primeira parte de uma formulação condicional que possui relevância interacional, isto é, à qual o receptor demonstrou orientação como um *evento de uso*.⁶⁷ A evidência é a de que B, diante dessa contingência e de seus propósitos imediatos, optou por fazer uso de um TRP fraco para assumir um turno pleno.

A meu ver, o favorecimento de estruturas como as sentenças em vários trabalhos da AC tem resultado em análises em que se nota claramente uma supervalorização desses esquemas gramaticais complexos em detrimento do processo de produção localizado da fala-em-interação. Lerner (1996), por exemplo, analisa o fenômeno da *finalização antecipada* no contexto de construções que envolvem mais de uma parte, tais como a condicional acima discutida ou ainda construções do tipo “quando + então”.

⁶⁶ O capítulo 4 irá discutir a relação entre unidades prosódicas e unidades sintáticas na fala espontânea. Como será apontado, as evidências são a de que a relação “oração simples-unidade entoacional” represente empiricamente cerca de 50% das instâncias de fala, e que a maioria dos outros 50% sejam representados por agrupamentos prosódicos de unidades abaixo do nível da oração. Daí a razão de considerar como “extensa” (i.e. marcada) qualquer contribuição que ultrapasse o escopo de uma oração simples concomitante a uma unidade entoacional.

⁶⁷ O termo *evento de uso* aqui utilizado segue a definição de Langacker (2001: 144-151), para quem “each such event consists of a comprehensive conceptualization, comprising an expression's full contextual understanding, paired with an elaborate vocalization, in all its phonetic detail” (p. 144). Ainda para o autor, as unidades linguísticas emergem precisamente a partir da abstração de aspectos comuns a esses eventos. O conceito será discutido em maiores detalhes na seção 4.3.

O tipo de fenômeno pode ser ilustrado pela passagem abaixo, reproduzida dos dados trazidos pelo próprio autor (p. 241):

Marty: Now most machines don't record that slow.

So I'd wanna- when I make a tape,

→ Josh: be able tuh speed it up.

Marty: Yeah.

O fenômeno estudado por Lerner mostra uma orientação clara de ambos os participantes à construção de esquemas complexos. A evidência dessa orientação está no fato de, em contextos desse tipo, os receptores se anteciparem sistematicamente em relação aos falantes correntes e oferecerem, eles mesmos, uma finalização do esquema previamente projetado. No exemplo acima trazido, Marty mostra com um sinal de reconhecimento “Yeah”, que a antecipação de Josh foi bem sucedida – tal como revela a maioria dos exemplos trazidos por Lerner.

A ênfase sobre estruturas como a sentença, porém, conduz o autor a uma interpretação dos dados que me parece equivocada. Para Lerner, construções do tipo “se + então”, “quando + então”, etc são exemplos de “TCUs-compostas”, formadas por um “componente preliminar” que projeta não apenas o seu término possível, mas também um segundo componente, chamado “final”. O término do componente preliminar constituiria um “ponto reconhecível” na trajetória dessa TCU-composta particularmente suscetível a incursões dos então receptores, desde que essas participações sejam restritas a práticas que envolvam um papel primário de receptor, isto é, práticas que não envolvam a disputa da palavra, como é o caso da própria finalização antecipada. Esses pontos reconhecíveis devem ser distintos dos TRPs, que segundo o autor, emergem apenas ao término possível do componente final da TCU-composta.⁶⁸ Assim, o fenômeno da finalização antecipada revelaria o “caráter semi-permeável” das unidades gramaticais (no caso, as sentenças) na fala espontânea.

É interessante notar como Lerner dá conta desse fenômeno empiricamente atestável que é a finalização antecipada. Em primeiro lugar, o autor aborda as unidades mínimas de construção do discurso usando duas terminologias distintas: TCUs e

⁶⁸ Pode-se dizer que Selting (2000), apesar de dar maior atenção à prosódia, chega à conclusão semelhante quando propõe que um conjunto de agrupamentos prosódicos com continuidade sintática (discutido na seção 3.4 acima) devesse ser interpretado como uma única TCU.

“componentes”. Algumas TCUs seriam formadas por um único “componente”, outras, como as condicionais, por vários. Embora o autor não discuta a natureza desses “componentes”, os dados sugerem que eles sejam unidades lingüísticas (i.e. sintagmas e orações simples). Em segundo lugar, o autor trata os pontos em que a troca de turnos se mostra relevante, novamente, usando duas terminologias: TRPs e “pontos reconhecíveis”. O TRP corresponde exclusivamente ao possível término da apódose, isto é, ao ponto projetado pelos recursos léxico-sintáticos para a completude de todo o esquema gramatical complexo, *onde o falante-corrente espera* que a palavra seja transferida; já os pontos reconhecíveis ficam reservados aos termos do(s) componente(s) preliminar(es), possivelmente separados por unidades entoacionais, onde *o falante-corrente espera* que a palavra permaneça com ele.

Retomando o exemplo fabricado de interação entre A e B anteriormente discutido a partir da perspectiva de Lerner, então, seríamos obrigados a interpretar a entrada de B na seta como uma “interrupção” no curso da TCU-composta formulada por A, num ponto que não se mostrava relevante para transição. O exemplo hipotético sugere, porém, que o término da prótase constituía sim um ponto relevante para transição, *senão para o falante-corrente, certamente para o seu receptor*. Tal reflexão chama a atenção, portanto, para o fato de a indeterminação das transições efetivas entre falantes ocorrer, pelo menos em parte, pelas diferenças de conhecimento, perspectivas e propósitos práticos que cada um dos participantes traz para a interação.

Seguindo a linha de raciocínio que vem sendo favorecida neste trabalho, então, ofereço um modo alternativo de interpretar os mesmos dados oferecidos por Lerner. Na realização de esquemas complexos do tipo “se-então”, restrições possivelmente cognitivas e interacionais *pressionam* um participante a formular a sua fala localmente por meio de breves incrementos sucessivos, produzidos em pequenos agrupamentos identificáveis e segmentáveis prosodicamente.⁶⁹ Os possíveis pontos de completude

⁶⁹ É importante destacar que as restrições cognitivas e interacionais *pressionam*, e não *determinam* o falante a produzir os turnos de maneira mais fragmentada (i.e. em agrupamentos prosódicos formados por orações simples ou constituintes menores). Estudos de viés cognitivista (e.g. Croft, 1995, 2005) têm oferecido hipóteses sobre as circunstâncias especiais sob as quais construções complexas tendem a ser agrupadas sob um único agrupamento prosódico (ver discussão da seção 4.6). Porém, considerando-se que a pressão pela quebra de esquemas complexos em constituintes menores possa também possuir uma motivação interacional, pode-se questionar qual seria o trabalho interacional que as diferentes opções de agrupamento desses esquemas realizariam em cada contexto específico.

No caso do estudo de Lerner, a sua transcrição em muitos casos deixa o estatuto prosódico do

dessas TCUs se mostram especialmente propensos à negociação, isto é, à re-calibragem, ao ajuste e ao re-acabamento da fala em curso por ambos os participantes. Em se projetando um esquema complexo pela formulação de um enunciado do tipo “Se ... ,” o falante sinaliza ao receptor que pretende levar a cabo um turno possivelmente composto por mais de uma unidade mínima, diminuindo a força de todos os possíveis TRPs que antecederem o TRP final, vagamente projetado pelo esquema complexo associado a essa instância.

Uma prática de negociação documentada por Lerner que se mostra recorrente nesses contextos, então, envolve a *finalização antecipada* (i.e. “antecipada” no sentido de não ser projetada pela fala-em-curso) no “espaço do turno” de seu interlocutor. Trata-se de uma incursão num ponto em que, a despeito da relevância para a transição em face do novo evento de uso, a palavra estava projetada para permanecer com o falante-corrente. Não por acaso, tais incursões *tendem* a preservar o direito à palavra com o falante-corrente, restringindo-se a práticas que envolvam um papel primário de receptor tal como a finalização antecipada, ainda que a possibilidade de um turno pleno ser tomado não deva ser descartada.⁷⁰ O risco desse ato de aposta do “intruso” multiplica a

componente preliminar duvidoso, o que, a meu ver, compromete diretamente a análise sobre os TRPs e os trabalhos dos participantes em orientação à fala-em-curso (para uma crítica sobre a subespecificação e o caráter intuitivo da notação prosódica nos trabalhos clássicos da AC, ver Couper-Kuhlen e Selting, 1996: 40-45).

⁷⁰ Em meu exemplo hipotético, procurei sugerir que incursões ao término dos “componentes preliminares” de construções como condicionais podem não envolver o papel primário de receptor, tal como projetado pela fala-corrente – ainda que sejam *constrangidas* a envolver esse papel. Essa possibilidade alternativa pode ser vislumbrada na discussão de Lerner sobre o exemplo abaixo (p. 253). O dado é trazido num contexto em que, o autor afirma, a completude do componente preliminar “algumas vezes” é usada como “recurso” para uma incursão que não “responde ao turno corrente”.

- 1 A: I was just gonna say come out and come over here and talk this evening.
- 2 But if you're going out [you can't very well do that.
- 3 B: [Talk you mean get drunk, don't you.
- 4 A: What?

Nessa passagem, pode-se argumentar que a incursão do então-receptor seja justificada pela natureza da intervenção: uma ação de reparo, cuja iniciação legitimamente se dá em qualquer ponto de uma fala em curso (Pedro Garcez, comunicação pessoal). Mesmo assim, permanece a questão de por que os então-receptores, nos casos em que intervém de maneira “prematura” e sem “responder ao turno corrente”, o fazem rotineiramente em orientação ao término da prótase.

Desse modo, a proposta de que os “componentes preliminares” também constituam unidades mínimas e de que a completude desses componentes constitua TRPs de menor força continua me parecendo uma explicação mais abrangente, capaz de abarcar tanto as ocorrências sistemáticas descritas por Lerner quanto as aparentes exceções, sem a necessidade de recurso, nesse caso, a outras categorias analíticas como “componentes” e “pontos reconhecíveis” – ainda que tal proposta necessite ser verificada empiricamente no futuro.

força de seu efeito: nos casos mais bem sucedidos, o interlocutor pode demonstrar uma sintonia ótima com o falante, fortalecendo sobremaneira o estatuto de sua relação e propiciando uma interpretação positiva de sua prática, como um ato colaborativo; nos casos mais mal sucedidos, porém, ele pode expor diferenças marcantes de perspectiva em relação ao falante, enfraquecendo o estatuto de sua relação e, eventualmente, propiciando uma interpretação negativa de sua prática, como um possível ato interruptivo.

Assim, qualquer projeto de fala que envolva um turno com mais do que uma única unidade mínima, seja um esquema “se-então”, seja um esquema de “história”, sofrerá pressão do sistema para a sua execução. Quanto maior o projeto de fala, maiores serão as dificuldades para manter os sucessivos TRPs internos ao projeto incólumes à intervenção do então-receptor. Nesse sentido, a diferença entre esquemas do tipo “se-então” e esquemas do tipo “histórias” é de grau: quanto maiores os projetos de fala, maior o risco de não alcançar o seu final e, conseqüentemente, mais ostensivas deverão ser as suas práticas de licenciamento (e.g. no caso das histórias, uma ou mais unidades inteiras deverão ser devotadas para o pedido de ratificação do projeto pelo receptor, i.e., o prefácio; ao passo que no caso de condicionais, a simples utilização de recursos léxico-sintáticos do tipo “Se + subjuntivo” parece bastar).

Na literatura da AC, a idéia de que a relevância para a transição pode ter forças diferentes dependendo do ponto na fala em curso no qual ela se manifesta foi apontada por Local e Kelly (1986) – até onde eu saiba, o único estudo que argumenta explicitamente a favor dessa idéia, ainda que de passagem.⁷¹ Ao analisar as características fonéticas de breves momentos de silêncio durante a fala, os autores identificam duas práticas distintas: os silêncios que sistematicamente retêm os turnos (realizados com fechamento glotal) e os que sistematicamente cedem o turno (não acompanhados de gestos articulatórios). Apesar dessa análise, os autores observam que,

⁷¹ Auer (2005) também aborda a questão de diferentes forças de projeção na interação, porém num sentido distinto do que eu aqui discuto. O autor se refere à previsibilidade com relação à natureza do elemento projetado, e não à questão da troca de falantes. Por exemplo, ele aponta que, no nível interacional, uma dada ação social pode antecipar *maximamente* uma segunda ação social (e.g. um chamado prevê fortemente uma resposta) ou *minimamente* (e.g. um “posso interromper você?” abre espaço para uma ação social a princípio não-especificada). Similarmente, no nível lingüístico, o uso de um elemento gramatical pode antecipar *maximamente* um outro elemento gramatical (e.g. uma preposição prevê a ocorrência de um substantivo no português) ou *minimamente* (e.g. um conectivo como “e” prevê uma continuação pouco ou nada especificada).

num número restrito de casos, os silêncios retentores de turno são pontos em que o então-receptor acaba iniciando um turno, embora tais incursões tendam a ser rapidamente abandonadas quando seguidas de sobreposição ou apresentem características de fala que legitimamente podem ocorrer em sobreposição. Tais observações levam os autores a concluir o que vem sendo argumentado neste trabalho: “Suspeitamos que o que está nos ocupando aqui são localizações na fala onde as reivindicações (*claims*) pelo estatuto de falante variam entre fracas (ou fortes) e que, embora os ‘silêncios retentores’ de fato retenham o turno de modo rotineiro e não problemático, eles não o fazem de maneira mecânica ou garantida” (p. 198).⁷²

Embora não expressa de maneira explícita, essa visão também está presente, a meu ver, no trabalho de Ford e Thompson (1996) sobre os “complexos pontos de relevância para transição” (CTRPs), nos quais converge a completude sintática, prosódica e pragmática. A idéia de que a conversa envolva, de um lado, TRPs que empiricamente se mostram menos preditivos para a transição de falantes, e, de outro lado, CTRPs altamente preditivos para essa transição, sugere que os pontos de possível transição ao longo da fala devam apresentar diferentes graus de relevância. Além disso, as autoras propõem que a completude pragmática, relativa à ação social em potencial realizada pela fala, pode se manifestar em dois sentidos: quando uma unidade constitui uma ação mínima dentro de ações sociais maiores como histórias; e quando uma unidade mínima não projeta nada além de si própria. Tal proposta abre espaço para a diferenciação do grau de força dos pontos de relevância para transição dependendo do contexto interacional em que emergem.⁷³

Assim, há razões suficientes, empíricas e teóricas, para se assumir que os TRPs se sucedem ao término possível de cada unidade mínima do discurso. Essas unidades mínimas fundamentais, contudo, não existem aprioristicamente, ou melhor, não têm sua

⁷² Essa proposta acrescenta à presente argumentação o fato de que o grau de relevância dos TRPs pode ser maior ou menor em decorrência não apenas do momento específico dentro da trajetória da atividade em curso, mas também do modo como os turnos-na-fala são construídos momento a momento, por exemplo, por meio da exploração de recursos fonéticos.

⁷³ As autoras reconhecem que o conceito de completude pragmática é bastante intuitivo e carece de maior definição, mas, apesar disso, entendo que a sua proposta esteja no caminho correto. A idéia de que uma ação pragmática mínima deva trazer como critério o fechamento de um contorno prosódico é exatamente o que tenho procurado argumentar neste trabalho, ao afirmar que os pedaços de fala encapsulados por unidades prosódicas devam ser entendidos como *eventos de fala* mínimos, ou *unidades de formulação* mínimas; isto é, ações sociais ou partes de ações sociais que se constituem como contingências de fala em si passíveis de questionamento, explicação, justificativa, etc.

relevância determinada aprioristicamente. Elas são *localmente formuladas* por meio do mapeamento de agrupamentos prosódicos sobre determinadas opções léxico-sintáticas que o falante seleciona em face de seus propósitos imediatos. A sintaxe, em si, oferece possibilidades a princípio infinitas de estruturação da fala, tendo em vista a sua natureza recursiva. É a prosódia, portanto, que contextualiza essa estrutura, transformando potenciais distribucionais em distribuições específicas e atribuindo a cada elemento um sentido preciso no todo, delimitado no aqui-agora diante de restrições cognitivas e interacionais.

Isso não impede que tanto o falante quanto o receptor monitorem a progressão da fala no decorrer da realização dessas unidades entoacionais, muitas vezes reagindo diante de fragmentos léxico-sintáticos que ainda não alcançaram completude prosódica. Parece haver evidências suficientes de que eles de fato o façam (e.g. Auer, 1996).⁷⁴

⁷⁴ Segundo Auer, a proposta de que a prosódia atue como um filtro entre a sintaxe e o sistema de tomada de turnos, especificando aos participantes quais possíveis pontos de completude sintática devem ser ouvidos como possíveis TRPs, não se sustenta nos seus dados. A evidência, para o autor, são as incursões comuns dos então-receptores em pontos da fala que coincidem com pontos de completude sintática (definida como um *gestalt* que não projeta nada além de si próprio), mas não prosódica (definida como o final identificável de um contorno entoacional).

Nos termos em que venho argumentando, a prosódia faz sim a conexão entre a sintaxe e a interação na medida em que ela determina o que, *com base nos propósitos práticos do falante*, deve ser tomado pelo receptor como uma unidade relevante interacionalmente. A sintaxe, em si, não constitui unidades, mas apenas potenciais de unidades. Quais desses potenciais serão de fato instanciados e, conseqüentemente *propostos* como relevantes interacionalmente, depende desse viés de quem produz, de quem formula. Se o receptor em muitos casos antevê a parte final dessas unidades mínimas de formulação, reagindo “prematuramente” (em relação às projeções da fala-em-curso) com base no monitoramento dos recursos léxico-sintáticos, tal ação é feita a seu total risco, de ter confirmada ou desconfirmada a sua antecipação (com as conseqüências interacionais que cada uma dessas opções acarreta). Desse modo, não apenas a incursão imediata do então-receptor deve ser levada em consideração para essa análise, mas a reação subsequente do próprio falante em face dessa incursão.

Um exemplo hipotético talvez seja ilustrativo desse argumento, que envolve, novamente, a necessidade de distinção entre a perspectiva do produtor e a perspectiva do receptor na interação. Goodwin (2002: 32), retomando a sua análise do artigo de 1979, argumenta que John, ao formular a sua fala por meio de um agrupamento prosódico que englobou “I gave up smoking cigarettes” estava realizando uma ação social bastante específica; distinta, por exemplo, de uma formulação que parasse em outro possível ponto de completude sintática, como “I gave up smoking”. O fato de ter ido até *cigarettes* tem conseqüência para a sua fala (e.g. abrir espaço para a possibilidade de ele continuar fumando outras coisas, em especial considerando o contexto dos anos 70, em que a conversa foi produzida).

Estendendo o argumento de Goodwin, consideremos ainda uma terceira formulação, constituída por dois incrementos distintos, “I gave up smoking. (breve pausa) Cigarettes”. Tanto na fala original de John, quanto neste caso hipotético, seria correto dizer que a transitividade do verbo *smoking* é especificada pelo elemento instanciado a seguir. Nesse último caso, contudo, essa transitividade é especificada a posteriori; isto é, somente após a conclusão de uma primeira formulação subespecificada é que o falante opta por remediá-la, expandindo o esquema com o elemento *cigarettes*.

O que o presente argumento propõe, então, é inverter a assimetria que tem sido tradicionalmente estabelecida entre o papel da sintaxe e o papel da prosódia na projeção local de pontos de relevância para a troca de turnos (envolvendo ou não transferência da palavra), agora em favor da prosódia. Tal hipótese parece encontrar uma sustentação empírica inicial (e.g. Ford e Thompson, 1996; Ono e Thompson, 1995),⁷⁵ embora o cerne dessa proposta seja de fato a tese – a ser aprofundada no capítulo 4 – de que os agrupamentos prosódicos expressam as unidades mínimas de formulação do discurso e que, dessa maneira, funcionem como *janelas de atenção* (Chafe, 1988) ou *quadros atencionais* (Langacker, 2001) básicos para a coordenação de ações sociais.

Outra questão a ser destacada sobre o papel desses recursos é a distinção entre projeções locais de TRPs e projeções distantes. Ao mesmo tempo que tendo a privilegiar a prosódia como recurso prioritário para projeção *local* de TRPs, tal recurso parece desempenhar um papel menos determinante no trabalho de projeções mais amplas, quando comparado, por exemplo, com os recursos lexicais e sintáticos (Auer, 1996; Selting, 2000) ou mesmo a gestualidade (Goodwin, 2002). Além disso, cabe ressaltar, como argumenta Goodwin (2002), que “[o] uso de múltiplos meios para construir ações oferece aos participantes a capacidade crucial de modificar as projeções das fronteiras de qualquer sistema isolado” (p. 48). Essa idéia implica, em consonância com a conclusão de Ford et al. (1996), que a fala situada envolve *múltiplas trajetórias*

Nesse último exemplo, a produção do segundo agrupamento revela claramente a orientação do próprio falante, de acordo com seus propósitos práticos, em relação aos possíveis problemas da subespecificação do verbo *smoking*. Esse novo incremento de formulação poderia ter sido motivado por uma mudança sutil de expressão ou de olhar de seu interlocutor, por exemplo, ou poderia ter sido motivado simplesmente por uma auto-reflexão e o desejo de não ter o seu comportamento reportado (parar de fumar) interpretado como o de uma pessoa que está ficando “careta”. Seja qual for o caso, a formulação da “mesma idéia” em um ou dois incrementos apresenta significados (e, conseqüentemente, possíveis repercussões interacionais) claramente distintas. Cada incremento é uma contingência a mais na interação e o falante torna-se responsável (*accountable*) por cada uma dessas formulações independentemente.

A reflexão acima, embora careça de subsídios empíricos para sustentá-la, me leva à consideração preliminar de que é problemática a proposta de SSJ, segundo a qual os sucessivos pontos demarcados por constituintes sintáticos (e.g. “I gave up [] smoking [] cigarettes []”) seriam todos TRPs independentemente do modo particular de instanciação dessas estruturas. Ainda que a prosódia não seja o único recurso em operação nesse trabalho de identificação de unidades mínimas, parece-me que, *pelo menos num nível local*, ela deva ter prioridade sobre a sintaxe.

⁷⁵ Ver também uma reconsideração que Jefferson (1986: 180-182) faz sobre suas análises do fenômeno da sobreposição, em vista das crescentes descobertas de lingüistas sobre o papel prosódia em contextos de fala situada (e.g. Local e Kelly, 1986).

que se inter-relacionam de maneiras sutis e complexas para produzir a coordenação da troca de turnos.

Ainda que essa pareça ser de fato a proposta mais robusta para dar conta do fenômeno da projetabilidade na conversação cotidiana em toda a sua complexidade, tal afirmação não é a explicação do fato; apenas o fato a ser explicado. Permanece em aberto a explicitação *do modo como* esses vários recursos e seqüências interacionais se inter-relacionam de maneira estruturada, organizada, e possivelmente hierarquizada, para produzirem diferentes TRPs, em orientação aos quais os receptores podem exercer seu direito pela palavra e/ou turno de maneira legítima.

O avanço dos estudos na AC e na lingüística interacional nos coloca hoje numa posição em que podemos começar a considerar mais precisamente essas “inter-relações sutis e complexas”, por exemplo, identificando sob quais circunstâncias alguns meios adquirem prioridade sobre outros; e considerando em que medida as “mudanças de projeções” acarretadas por essas inter-relações deveriam ser interpretadas em termos absolutos (i.e. de bloqueio, suspensão ou eliminação de TRPs em contextos locais, como vários autores têm proposto) ou em termos mais relativos (i.e. de relativização do grau de relevância de um ou mais TRPs estruturalmente previstos, como o presente trabalho propõe).

3.6. Estrutura, contingência e as unidades mínimas do discurso

Diante de uma visão um tanto rígida dos conceitos de turno, unidades mínimas e TRP, é comum encontrarmos na literatura afirmações como as de Selting (2000), segundo a qual “é impossível fazer qualquer julgamento sobre o possível término de unidades [na interação] de uma forma livre de contexto” (p. 488); ou conclusões como a de Ford et al. (1996), segundo as quais a busca por definição das unidades do discurso pode nos conduzir a um caminho equivocado, nos impedindo de enxergar as práticas que são empregadas pelos participantes e que permitem aos seus interlocutores (e aos próprios analistas) interpretar os seus turnos (p. 450).

Embora o argumento por trás dessas declarações seja válido (i.e. uma visão sobre gramática que enfatiza a sensibilidade da língua às contingências da situação), tal afirmação parece supervalorizar a distinção entre língua e discurso. Será, de fato, impossível fazer *qualquer julgamento* sobre trajetórias de completude de unidades de

uma forma livre de contexto? Seria, como se sugere, indesejável haver análises voltadas para a definição das unidades da interação? Ou não seria mais cogente dizer que julgamentos possíveis sobre as trajetórias de unidades num nível livre de contexto devam se mostrar suficientemente abertos e flexíveis para dar conta da contingencialidade da fala?

A meu ver, a orientação às unidades mínimas do discurso e aos TRPs num nível livre de contexto é não apenas possível, mas desejável. É o que permite, por exemplo, aos participantes *projetar* o término de uma dada ação tão logo ela se inicie, ainda que esse ato de projeção tenha que ser flexível o suficiente para, eventualmente, precisar mudar o seu curso ou mesmo ser abandonado e refeito em meio à trajetória da fala. A projeção é, portanto, uma ação ao mesmo tempo livre de contexto e sensível ao contexto. Ela é livre na medida em que a instanciação inicial de certas práticas convencionalizadas permite uma antecipação aproximada do curso completo que a interação imediata supostamente deverá tomar, o que é crucial para a tomada de decisões dos participantes, seja para submeter-se ou não ao projeto antevisto; mas ela é também sensível ao contexto na medida em que a projeção não garante o resultado final, pelo fato de o monitoramento mútuo dos interlocutores momento-a-momento ser capaz de suscitar reorientações em face de contingências que não estavam previstas.

Esse é precisamente o jogo entre a estrutura e a contingência que está na essência da conversação: sem a primeira dimensão, os falantes passariam por maus bocados para a realização de ações sociais que demandam coordenação – vitais para a solução de problemas práticos na experiência cotidiana de seres sociais como os de nossa espécie; sem a segunda dimensão, a dificuldade seria a de tornar essas mesmas ações sociais coordenadas bem sucedidas diante dos novos desafios inevitavelmente postos por novas experiências. Da primeira dimensão, tiramos proveito da estabilidade e da economia cognitiva que ela acarreta para resolver os problemas mais recorrentes em nosso dia-a-dia; da segunda, alcançamos a flexibilidade cognitiva e social necessária para lidar com a novidade ao mesmo tempo em que enriquecemos o nosso repertório de ferramentas para enfrentar desafios futuros.

A grande elegância do sistema de tomada de turnos proposto por SSJ reside exatamente nessa dupla propriedade, de ser ao mesmo tempo livre-de-contexto e sensível ao contexto. Os TRPs ao longo da cadeia de fala, por exemplo, são estruturalmente previstos pela organização convencional da língua, embora esses

mesmos TRPs sejam diretamente afetados (i.e. enfraquecidos, fortalecidos, aproveitados, passados, etc) dependendo das estratégias que os participantes levam a cabo em situações específicas de fala em tempo real. A relação entre gramática e interação, no final das contas, consiste num caminho de mão dupla. A estrutura organiza as formas de participação na interação; enquanto a contingência reorganiza a estrutura imanente numa fala específica e, conseqüentemente, o próprio sistema abstrato que subjaz a essa fala (Schegloff, 1996).

Apesar disso, no que diz respeito às unidades mínimas do discurso, observa-se na literatura da AC, desde os trabalhos clássicos, uma maior valorização da contingência do que da estrutura. Essa visão se reflete, a meu ver, numa reafirmação ostensiva de que a unidade de construção de turnos *não* seja uma unidade de natureza lingüística, mas sim interacional. Como Schegloff (1996) aponta, o termo TCU é usado para se referir a “unidades que *podem* constituir turnos possivelmente completos” em que a “transição para o falante seguinte se torna relevante (mas não necessariamente realizada)” (p. 55, itálicos no original). Retomando o seu argumento, o autor enfatiza que não se trata de uma discussão terminológica, isto é, de substituir TCU por sintagma, oração, sentença. Trata-se de enfatizar que a fala que habita os turnos é produzida em tempo real e sujeita às contingência interacionais, de modo que os “objetos sintáticos *estáticos* de grande parte da teorização lingüística” devem se mostrar “diferentes em *ênfase*” e “talvez em *tipo*” das unidades encontradas na fala-na-interação (p. 56, itálicos meus).

Tradicionalmente, essa visão sobre a gramática de fato se mostrou incompatível com as diferentes correntes de teorização lingüística – uma exceção notável sendo a filosofia da linguagem de Bakhtin (1979, publicado originalmente em 1929). Mais recentemente, porém, novas teorias lingüísticas têm oferecido um modo alternativo de enxergarmos a língua e o seu modo de funcionamento, além de oferecer uma caracterização formal e funcional bem mais precisa sobre as unidades mínimas do discurso, diferentemente do que se observa com o termo TCU.

Uma dessas teorias é a da *gramática cognitiva* de Langacker (1987; 1991; 2000; 2001). Para o autor, os tipos de constituintes que aparecem nas análises sintáticas clássicas não são vistos como essenciais para a estrutura da língua; mas sim como emergentes a partir de instâncias concretas de uso. Por exemplo, em eventos de fala reais, pode-se dizer que a ocorrência de orações como “A gente deu uma bola pro

menino”, “Ele deu o dinheiro pra Maria”, “Eu dei gorjeta pro garçom”, e incontáveis outras ocorrências análogas, permite a emergência de esquemas bitransitivos altamente esquemáticos do tipo “NP + VP + NP + PP”. Contudo, esse alto grau de esquematização não seria atingido sem a emergência prévia de esquemas em níveis de representação mais baixos, isto é, mais dependentes do contexto (e.g. NP + deu + NP + pra + NP). Entre a instância concreta e o esquema bitransitivo mais abstrato, há um longo contínuo de representações com diferentes níveis de esquematicidade disponíveis aos falantes para a construção de suas formulações. Langacker (2000: 125) argumenta, nesse sentido, que as variantes mais dependentes do contexto, mais concretas, devem ser mais fundamentais para o uso da língua do que os esquemas abstraídos de todo contexto – uma abordagem de baixo-para-cima que contrasta fundamentalmente com a abordagem de cima-para-baixo das vertentes majoritárias na lingüística.

Outro trabalho que caminha nesse sentido é o de Hopper (2001) e a sua proposta da *gramática emergente*. Para o autor, “as construções gramaticais são racionalizações normatizadas de famílias de partes quase-lexicais fragmentadas e menores. Elas resultam da gramaticalização desses fragmentos” (p. 124). Entre as evidências empíricas desse favorecimento de instâncias mais concretas (partes quase-lexicais) sobre as mais abstratas (construções gramaticais), encontra-se a constatação, baseada em corpus de uso espontâneo da língua, de que vários tipos de construções tendem a estar “presos” a determinados itens lexicais.⁷⁶

Por exemplo, na análise do autor sobre as pseudo-clivadas, Hopper verifica que os verbos nessas estruturas provêm de um conjunto lexical bastante restrito no inglês (*do, happen, need, make, use, say, tell* e alguns outros).⁷⁷ Essa natureza quase-lexical

⁷⁶ Tomasello (1992a) e Tomasello et al. (1997) mostram que a dependência de esquemas em relação a itens lexicais se manifesta desde os primeiros anos de aquisição de língua, a criança mostrando uma produtividade limitada, exceto no que diz respeito à variação de “estruturas baseadas em itens” com espaços a serem preenchidos altamente restritos.

⁷⁷ Essa observação leva o autor a questionar a proposta predominante de que a pseudo-clivada tenha a função principal de salientar o constituinte predicador (que carregaria a informação nova) e colocar em segundo plano o constituinte da oração –wh (que carregaria a informação dada). Ao observar a natureza dos verbos que compõem o pequeno conjunto das pseudo-clivadas (verbos semanticamente simples, básicos e gerais), Hopper propõe que a função central dessas construções seja a de protelar uma predicação específica a ser realizada a seguir (e.g. *do* projetaria uma ação específica; *happen*, um acontecimento específico, etc), por razões pragmáticas tais como a de introduzir gradativamente um tema complexo, ganhar tempo para o planejamento da fala, entre outras. A salientação do material contido no predicado, nesse sentido, seria então um epifenômeno dessa função de protelação.

das construções gramaticais mostra que, no processo de produção e compreensão da língua, o recurso a esquemas prototípicos altamente abstratos em face de instâncias concretas da fala não deva ter um papel central; diferentemente, os próprios enunciados devem ser entendidos como reflexos de estruturas imanentes. Com base nessas reflexões, o autor conclui:

I suggest that the demands of consistency, the requirement that grammatical constructions be measured against a single ideal prototype, lead us to privilege in linguistic analysis sentence types that have achieved this consistency only in the context of reflective, planned discourse, and that in natural spoken discourse the visible parts of these canonical construction types appear as fragments. But these fragments should not be seen as incomplete partials. Instead, the reverse is the case: the fuller ‘canonical’ constructions should rather be seen as highly stylized cultural artifacts, amalgamations of fragments put together and grammaticalized through stylistic and normative conventions (p. 125).

Como a discussão acima mostra, propostas como as de Hopper e Langacker sobre a gramática revelam a mesma preocupação de Schegloff (1996) em relação à relevância das construções gramaticais clássicas nas análises do discurso espontâneo, tanto em termos de “ênfase” quanto de “tipo”. Mais do que isso, a visão lingüística trazida por esses pesquisadores (ver também Bybee, 2006 e Tomasello, 2000a, entre outros trabalhos da chamada “gramática baseada no uso”) permite, a meu ver, uma caracterização formal das unidades do discurso que se mostra teoricamente rigorosa e empiricamente fundamentada; ao contrário da caracterização um tanto intuitiva (ou mesmo da total ausência de definição) das unidades mínimas do discurso na maioria dos trabalhos da AC. No capítulo 4, então, a caracterização lingüística dessas unidades mínimas da fala espontânea será discutida em maior profundidade.

Como argumentam Ford et al. (1996), a busca pela estrutura imanente nas unidades do discurso pode *sim* nos conduzir a um caminho equivocado, mas não necessariamente; tudo irá depender do modo como a contingência será tratada: como uma imperfeição da fala espontânea ou como um dos seus principais valores (Schegloff, 1991a). E se a abstenção dessa busca por unidades parece viável para os pesquisadores de línguas como o inglês, isso ocorre tão somente em decorrência do amplo conhecimento lingüístico acumulado sobre o qual tais investigadores podem se apoiar. No caso das LSs, diferentemente, tal empreendimento é não apenas desejável como também indispensável, em especial quando assumimos, como na atual pesquisa,

premissas teóricas básicas estabelecidas pelo próprio campo da AC (cf. citação de Schegloff na seção 3.2).

3.7. Conclusão

O presente capítulo mostrou a relevância da linha teórica conhecida como AC para a abordagem desta pesquisa. O estudo sobre a segmentação da libras surgiu a partir da sugestão de um trabalho já antigo na área (Baker e Padden, 1978), que destacou a necessidade de adotar, no estudo das LSs, metodologias que minimizassem a influência da estrutura da língua oral sobre os dados sinalizados; e que trazia como sugestão, com base nessa preocupação, a proficuidade de se observar a troca de turnos em conversas entre surdos com o intuito de se buscar pistas para uma segmentação menos enviesada das LSs.

Paralelamente a esse trabalho, a AC vinha demonstrando, começando pelo trabalho seminal de SSJ e se estendendo nas décadas subseqüentes com a contribuição de lingüistas, que a estrutura da língua está intimamente relacionada à estrutura das atividades sociais que lhe dão vida. No caso da conversação, a organização da troca de turnos é um dos aspectos estruturados dessa atividade que aponta reflexivamente para a organização gramatical das línguas. Com base nisso, o projeto de pesquisa assumiu que alguns conceitos operacionais da AC e o conhecimento acumulado na área ofereceriam importantes subsídios teóricos e metodológicos para esse desafio de encontrar um critério sólido de segmentação do discurso sinalizado em unidades gramaticais.

Apesar disso, um aspecto particularmente importante da teoria revelou uma limitação: a caracterização das unidades de construção de turnos. Nesse aspecto, surge um curioso paradoxo: ao mesmo tempo em que os estudos da área argumentam a favor da natureza distinta das TCUs em relação à língua, enfatizando a sensibilidade dessas unidades ao contexto, nota-se nesses mesmos estudos uma visão bastante tradicional da lingüística, pelo modo como a corporalidade da fala, manifesta prosódica e gestualmente, é colocada em segundo plano nas análises em favor do léxico e da sintaxe.

Tradicionalmente, os estudos lingüísticos revelaram um esforço de retratar como incompatível a relação entre língua e discurso, talvez por uma necessidade política de delimitação do campo em relação a outras ciências humanas como a antropologia, a sociologia e a história. Contudo, visões mais recentes de língua têm apontado mudanças

importantes nesse modo de olhar, se colocando como uma alternativa interessante para dois dos problemas acima apontados: a questão da sensibilidade da língua aos contextos de uso e a da inclusão da corporalidade da língua nas teorizações sobre a gramática. Neste capítulo, a primeira questão foi brevemente abordada, na discussão da relação entre o uso e o sistema abstrato da língua; no capítulo 4, a seguir, a segunda questão será aprofundada na discussão sobre o papel que a prosódia e a gestualidade desempenham no gerenciamento da atenção e as implicações disso para a questão da segmentação das línguas em unidades gramaticais.

4. Fundamentação teórica II: Língua e discurso na perspectiva da gramática baseada no uso

4.1. Introdução

Ainda que o conhecimento e/ou adoção dos princípios teóricos e metodológicos da AC por lingüistas seja restrita (e.g. Selting e Couper-Kuhlen, 2001 e o campo da recente lingüística interacional), várias propostas de investigação da língua têm revelado uma afinidade e uma abertura teóricas suficientes para que articulações entre ambas as áreas possam ser buscadas. É o caso, por exemplo, da lingüística funcional e da lingüística cognitiva. Em ambas as áreas, os pesquisadores partem do pressuposto de que o discurso tenha um papel determinante na emergência, consolidação e evolução da língua. No funcionalismo, essa premissa tem se revelado na análise de vários fenômenos gramaticais clássicos tomando-se por base dados empíricos provenientes do discurso espontâneo (e.g. a questão da transitividade, em Hopper e Thompson, 1980; de tópico, em Givón, 1983; de estrutura da informação, em Chafe, 1994). Já na lingüística cognitiva, essa mesma premissa tem se revelado mais fortemente em sua própria formulação teórica, de acordo com a qual a organização da língua não pode ser entendida como desvinculada ou autônoma em relação a outros domínios da experiência humana, como a cognição e a interação (Langacker, 1987; Lakoff, 1987; Johnson, 1987; Fauconnier, 1985).⁷⁸

Como discutido no capítulo anterior, o presente estudo traz uma perspectiva sobre a relação entre gramática e interação fortemente influenciada pela AC, embora, no que diz respeito à utilização de certos conceitos operacionais tais como as unidades de construção de turnos, a abordagem sociológica se apresente limitada para fundamentar uma proposta de segmentação da libras. A aposta teórica desta pesquisa, então, é a de que essa limitação possa ser remediada considerando-se trabalhos provenientes dessas duas áreas afins da lingüística: com orientação mais funcional, os estudos que se voltaram para a investigação das unidades básicas da interação espontânea (e.g. Chafe,

⁷⁸ Apesar dessa orientação teórica, a lingüística cognitiva tem se desenvolvido predominantemente a partir da análise de dados fabricados e/ou experimentais. Por esse motivo, mais recentemente alguns pesquisadores têm procurado desenvolver essa linha tanto teórica quanto metodologicamente a partir da devida consideração de dados provenientes do discurso espontâneo (e.g. Hougaard, 2005; Sinha, 2005; Tomasello, 2000a).

1994; Croft, 2005; DuBois, 1987); e, com orientação mais cognitivista, os trabalhos que se voltaram para a investigação das manifestações do fenômeno de figura-e-fundo na língua (Langacker, 1987) e a sua relação estreita com o estabelecimento da atenção conjunta (Langacker, 2001; Tomasello, 2000a).

Neste capítulo, então, será discutido o conceito-chave de *unidade entoacional* (UE) e algumas das possíveis restrições que esse conceito revela sobre o processo de produção da fala espontânea. Serão discutidas as características, a função e o potencial universal dessas unidades, tendo em vista a relevância do conceito para o trabalho de segmentação do discurso. Com relação à função dessas unidades, será ainda discutido de que modo as UEs podem ser entendidas como uma das manifestações lingüísticas de um fenômeno prevalescente na organização da experiência humana: a *salientação*, relativa ao estabelecimento de relações de *figura-e-fundo*.

Nesse ponto, então, o desenvolvimento da discussão permitirá que nos reconectemos com a discussão central do capítulo anterior, sobre o fenômeno da projetabilidade e da coordenação da ação social na interação. Ocorre que os recursos de salientação se mostram fundamentais não apenas para a organização da experiência perceptual e lingüística, mas também da experiência social, viabilizando o estabelecimento da *atenção conjunta*, um dos pilares fundamentais do desenvolvimento ontológico da língua e cultura humanas.

Da mesma maneira que, a meu ver, o campo da AC pode se beneficiar da perspectiva oferecida por essas propostas, entendo que o olhar da AC deverá também trazer importantes questionamentos a respeito de determinados pressupostos reducionistas sobre a relação entre o individual e o social, entre a gramática e o discurso, que se mostram ainda bastante arraigados nessas vertentes da lingüística. Nesse sentido, o capítulo anterior se colocará como importante referencial teórico para uma leitura crítica das propostas que serão aqui articuladas, e as devidas reconsiderações deverão ser introduzidas ao longo da exposição teórica.

4.2. As unidades entoacionais

Da mesma forma que a observação descompromissada de uma conversação revela, grosso modo, que os interlocutores procedem por meio da troca de turnos de fala, a observação da fala individual – seja em turnos de fala conversacional, seja em contextos menos dialógicos tais como o de narrativas, palestras, etc – revela, por sua

vez, que o discurso procede de maneira intermitente, por meio de sucessivos agrupamentos de atividade verbal. E da mesma forma que os estudos da interação têm procurado investigar a natureza e função desses turnos na fala, estudos lingüísticos têm se dedicado, há pelo menos quatro décadas, à investigação desses agrupamentos básicos da língua, particularmente delimitáveis por meio da prosódia (e.g. Halliday, 1967; Crystal, 1969; Pierrehumbert, 1980; Selkirk, 1984; Brazil, 1985; Cruttenden, 1986; Nespore e Vogel, 1986; DuBois et al., 1992; Chafe, 1994).⁷⁹

Para Chafe (1994), a intermitência característica do discurso pode ser entendida a princípio como produto de uma necessidade biológica. Tendo em vista que o ato de falar exige que exalemos ar dos pulmões, o uso da língua precisa envolver períodos de suspensão da vocalização para que o ar seja inspirado em meio ao fluxo da fala. O autor argumenta, então, que essa necessidade fisiológica opera não de maneira aleatória, mas em harmonia com os segmentos funcionais básicos do discurso – que ele chama de *unidades entoacionais*.

Tais unidades não são definidas exclusivamente em termos de contornos entoacionais coesivos, ainda que esse seja um dos principais critérios para a sua delimitação (DuBois et al., 1993). Entram em jogo também a presença ou ausência de vocalização (percebida como pausa), mudanças na frequência fundamental (percebida como pitch), mudanças na duração (percebida como encurtamento ou alongamento de sílabas), mudanças na intensidade (percebida como altura) e mudanças na qualidade da voz de várias formas (Chafe: 1994: 58). Além disso, Chafe aponta também um último critério, decorrente da natureza social da interação: as mudanças de turno.⁸⁰

⁷⁹ Dentro da perspectiva interacional que o presente estudo adota, é importante assinalar, como Ford e Thompson (1996: 145-8), que os estudos em questão se diferenciam em pelo menos dois grupos por suas abordagens teóricas e metodológicas. Dentre as propostas mais afins à presente pesquisa, estão os estudos baseados em dados perceptuais-auditivos como os de Chafe, DuBois, entre outros, que se desenvolveram a partir da análise de conversas e narrativas espontâneas ou quase-espontâneas (i.e. eliciadas); do outro lado estão os estudos baseados em dados acústicos, como os de Nespore e Vogel, Selkirk, entre outros, que se desenvolveram a partir da análise de sentenças escritas lidas em voz alta em contextos imaginados. Nesse último caso, a diferença mais marcante decorre do fato de construções sintáticas serem assumidas aprioristicamente, o trabalho do investigador sendo o de analisar como os agrupamentos prosódicos mapeiam essas unidades pré-estabelecidas. Estudos que buscam confrontar dados perceptuais e acústicos são raros, mas sugerem a proficuidade de contrastar análises perceptuais com acústicas (Schuetze-Coburn et al., 1991).

⁸⁰ Como a discussão acima revela, o termo “unidades entoacionais” pode envolver outros recursos prosódicos, além da entoação, para a sua delimitação. Embora essa observação sugira que o termo seja inadequado, opto por mantê-lo neste trabalho tendo em vista a sua ampla aceitação na literatura.

Analisando uma UE prototípica, Chafe aponta alguns importantes critérios analíticos formais para a sua delimitação. A ausência de vocalização frequentemente (mas não necessariamente) ocorre ao início e ao término de uma UE. Segundo Chafe, “uma precisão [na medição] de centésimos de segundo é mais do que adequada” (p. 59). Com relação à duração, o padrão aceleração-desaceleração das UEs, responsável por reduções fonológicas ao início da unidade e alongamentos ao final, pode ser tomado como o principal critério de delimitação de UEs. No que diz respeito ao pitch, a descrição de Chafe discrimina três tipos de contornos entoacionais: o terminal descendente, o terminal ascendente e o não-terminal ascendente. E sobre a qualidade da voz, destaca-se como critérios adicionais a presença comum da *creaky voice* no início e no final das UEs, entre outras possíveis mudanças nesse critério de uma unidade para outra.

A tese central de Chafe é a de que o fenômeno das UEs reflete restrições de processamento cognitivo dos seres humanos. Todos possuímos uma representação de mundo que envolve uma imensa gama de informações, adquiridas ao longo de nossa experiência. Esse vasto conhecimento, porém, não estaria disponível ao falante como um todo nas situações de interação; diferentemente, apenas uma pequena porção poderia ser focalizada de maneira ativa a cada novo momento. Chafe chama de *consciência*⁸¹ essa “focalização ativa de um ser consciente sobre uma pequena parte do seu modelo auto-centrado de mundo” (p. 28), e argumenta que as UEs sejam a manifestação lingüística desse fenômeno psicológico. Para ajudar a compreender o tipo de limitação envolvida na produção das UEs, então, Chafe evoca o funcionamento da visão, que segundo ele, opera de maneira similar: com o deslocamento contínuo do foco de uma região para outra ao longo do tempo, acompanhado de uma periferia visual, que permanece acessível mas sem uma orientação direta do observador, e de regiões mais amplas que se mantêm inativas ou inacessíveis à visão enquanto o foco não se desloca em direção a elas.

⁸¹ O conceito de “consciência” está entre os mais debatidos da psicologia e não é o objetivo deste trabalho assumir um ou outro sentido. Embora a proposta de Chafe seja definitivamente relevante para o meu trabalho, entendo que a maneira mais adequada de denominar esse “foco de ativação” de que o autor trata, seja pelo termo “consciência” ou qualquer outro, é irrelevante para a presente discussão. A utilização do termo “consciência” nesta seção, portanto, é mantida exclusivamente em referência à perspectiva de Chafe.

Essa tese sobre as restrições no processo de ativação de informações no uso da língua é elaborada a partir da análise das características dos pedaços de fala que aparecem correlacionados às UEs. Chafe observa que esses pedaços apresentam um número bastante restrito de palavras; que eles são formados em sua maioria por orações simples ou constituintes menores; e que eles introduzem uma quantidade restrita de informações a cada unidade.

Um outro modo de interpretar a tese central de Chafe, embora ele não coloque a questão explicitamente nestes termos, seria dizer que a prosódia carrega uma função gramatical básica de salientação de informações (Langacker, 2001). Tal visão permite unificar o papel da prosódia no âmbito do discurso e no âmbito dos constituintes que ela agrupa: no discurso, a UE como um todo refletiria um foco de ativação,⁸² que se impõe em relação à unidade anterior, ao passo que, dentro da própria UE, recursos prosódicos serviriam para salientar certos elementos em detrimento de outros.

Com relação a essa última dimensão, Chafe propõe uma análise mais localizada sobre o estatuto das informações dentro de uma UE, buscando atribuir a qualificação de informação “nova” ou “dada” de maneira independente a cada referente que participa num evento ou estado. Essa visão local permitiu ao autor evidenciar que a língua confere maior proeminência às idéias novas em relação às idéias dadas, e que entre esses dois níveis é possível identificar um terceiro nível, de saliência secundária (p. 71). Esses graus de proeminência se manifestariam lingüisticamente por meio de desvios na base do pitch para cima ou para baixo, opcionalmente acompanhados de aumento de volume e duração.

Reconsiderando os estudos anteriores voltados para esse mesmo tema, Chafe se contrapõe à noção de que a informação nova seja aquela que “aparece pela primeira vez no discurso”, ou que seja “a informação nova para o interlocutor”; e que informação dada seja aquela que “já apareceu no discurso”, ou que “já é conhecida pelo interlocutor”. Para Chafe, não é o discurso ou o conhecimento de mundo dos interlocutores que orienta o estatuto informacional dos elementos de uma UE, mas a *consciência* dos interlocutores, isto é, o *estado atual de ativação* de determinados itens informacionais na mente dos interlocutores. Como o falante somente tem acesso direto à

⁸² O termo “foco” e/ou “foco de atenção” neste trabalho está sendo usado num sentido amplo e permanece em aberto quais seriam as suas possíveis relações com o conceito de *foco* estudado na lingüística.

sua própria consciência, o modo como ele expressa informações numa UE vai depender, portanto, de uma avaliação sobre os estados de ativação de informações na mente de seu interlocutor, que podem variar entre ativo, semi-ativo e inativo. A crença de um interlocutor sobre esses estados teria várias fontes: o contexto circundante de interação, o discurso decorrente até então, as interações sociais prévias à atual e as experiências e culturas compartilhadas.

Desse modo, Chafe afirma que são expressas como “novas” as informações *ativadas pela primeira vez neste ponto da conversação*. Em contrapartida, informações “dadas” são aquelas consideradas *já ativas na mente do interlocutor neste ponto da conversação*. Além disso, o autor distingue um estado intermediário entre as duas, de informações “acessíveis”, caracterizadas como informações ativadas a partir de um estado semi-ativo. A cada um desses tipos de informação, portanto, será associado um diferente custo cognitivo de ativação: as informações dadas envolvem o menor custo de ativação; as acessíveis, um custo um pouco maior; e as novas, o maior esforço mental por envolverem a conversão de uma idéia inativa para o estado ativo.

Com relação à função dessas unidades, o autor propõe que

During these successive activations [of the speaker's focus of consciousness] the minds of the speaker and the listener are necessarily out of phase. At the completion of an intonation unit the speaker must intend that a reasonable facsimile of his or her focus of consciousness will have become active in one or more other minds. It is through this dynamic process of successive activations ... [t]hat language is able to provide an imperfect bridge between one mind and another (Chafe, 1994: 63).

Em outra oportunidade, Chafe (1988) aborda essa mesma função comunicativa com uma metáfora interessante. Ele afirma que as UEs “nos oferecem *janelas valiosas para o fluxo do pensamento* ao mostrar quantos e quais tipos de informação estão sendo focalizados em diferentes momentos, bem como (...) os tipos de movimento que ocorrem de um pedaço focalizado de informação para o outro enquanto o pensamento está sendo verbalizado” (p. 2-3, ênfases minhas).

A função comunicativa das UEs vai ser então, num segundo momento, especificada pelo autor na distinção entre três tipos de unidades: as substantivas, as regulatórias e as fragmentárias. As UEs substantivas seriam aquelas que trazem alguma idéia sobre um estado, evento e/ou referente, e constituiriam o tipo mais recorrente nas interações. As UEs regulatórias, similares aos *marcadores discursivos* (Schiffrin, 1987),

seriam aquelas que servem para regular a interação e o fluxo de informação, podendo ser subdivididas em unidades que Chafe classifica como “textuais” (e.g. *and then, well*), “interacionais” (e.g. *mhm, you know*), “cognitivas” (e.g. *let me see, oh*) e “validacionais” (e.g. *maybe, I think*). Para Chafe, essas unidades substantivas e regulatórias, consideradas “bem sucedidas” (p. 63), contrastam com as UEs fragmentárias ou truncadas, que não chegam a alcançar o final de seus contornos prosódicos.

Essa caracterização geral do trabalho de Chafe (1994) sobre as UEs revela que a sua proposta, embora relevante para os propósitos da presente abordagem de pesquisa, deva ser reinterpretada sob um prisma distinto. De fato, a constatação empírica de que o discurso espontâneo se segmenta em UEs, que agrupam pedaços de fala bastante restritos lexical e sintaticamente, exige do analista – assim como Chafe procura fazer – um questionamento sobre o porquê desse modo de estruturação do discurso. Nesse sentido, a análise de Chafe sobre a relação das UEs com o que ele chama de *consciência* (o que está, no momento da interação, sendo ou não focalizado pelos interlocutores), e não mais em termos de *conhecimento* (o que os interlocutores sabem ou não), desloca o foco de objetos mentais estáticos para o caráter seletivo, estratégico e dinâmico da fala no aqui-agora da interação espontânea.

Apesar disso, o modo como a proposta como um todo é articulada revela vários pressupostos mentalistas da psicologia cognitiva que não são capazes de explicar alguns fenômenos corriqueiros da interação. Como bem aponta Edwards (1998: 92), o trabalho retrata o participante de uma interação como um interlocutor ingênuo, que individualmente faz sentido das experiências mundanas sob a restrição de suas limitações de processamento de informação, e se coloca diante de um outro interlocutor a fim de receber ou transmitir experiências “interessantes”.⁸³ Assim, ao caracterizar a situação interativa como um espaço de transferência de informação e não como, primordialmente, um domínio de ação social, a proposta simplifica o trabalho psicológico em que os interlocutores estão envolvidos em suas interações cotidianas, o modelo de transferência sendo apenas uma das possibilidades de arranjo dos interlocutores para o gerenciamento da informação (Goffman, 1974).

⁸³ Goffman (1974) se refere a essa visão do interagente como uma super-racionalização do indivíduo (p. 515) e, em contraposição a ela, advoga a necessidade de enxergarmos o espaço de interação de maneira análoga a um palco de teatro – no sentido performático do termo.

Tomemos, por exemplo, a expressão de uma informação dada, ou nova, ou acessível numa dada UE, por meio da exploração de recursos prosódicos como pitch, volume e duração. Como argumenta Edwards (1998: 99), não é a análise lingüística de Chafe que está sendo questionada, mas a sua análise psicológica. Na conversação cotidiana, tais recursos são empregados pelos falantes estrategicamente, com fins retóricos específicos e de acordo com os seus propósitos, e não exatamente por meio de uma avaliação ingênua sobre o possível estatuto informacional de certos itens na mente de seus interlocutores. A reinterpretação de Edwards sobre um exemplo trazido por Chafe é reveladora desse argumento:

Take, for example, an utterance that crops up a number of times in [Chafe's book], 'I tálked to Lárry last nìght' (the accents mark emphasis, with primary stress on Larry). Chafe remarks that 'for the participants in the conversation ... there must have been one person named Larry who was more salient than any others possibly known to them by that name' (p. 101). This is a reasonable assumption, but a logical and normative one (rather than empirically grounded), based on the presumption of conversation as disinterested, doing-nothing information exchange. Imagine the utterance as a way of fishing for the other person to ask who Larry is, where he is perhaps someone important (their boss) or famous, like Laurence Olivier. Calling him Larry can be a way of signaling the speaker's casual acquaintance with someone so famous, where the casual mention of him is a way of 'doing being casual' about knowing people like that. And it can be a way of getting the basis of that acquaintance invited as a topic by the addressee, like with 'Larry who?', where it would be wrong to assume that 'Larry who?' signals some kind of communicative or psychological error about shared knowledge on the part of first speaker (p. 100).

O argumento de Edwards, como o próprio autor assinala, não envolve uma proposta de interpretação "preferida" sobre a passagem; diferentemente, trata-se de apontar uma interpretação que o modelo de comunicação como transmissão de informações, uma redução psicológica sobre o que os participantes *podem* fazer na interação, não é capaz de dar conta. Além disso, do ponto de vista metodológico, as hipóteses de Chafe sobre o estatuto informacional de itens da fala envolvem claras dificuldades analíticas, pois têm que ser buscadas fora dos dados, em especulações sobre o que se passa na mente dos interlocutores (Edwards, 1998: 99); ao passo que a

abordagem interacional privilegiaria o modo como a própria fala dos participantes poderia servir como evidência para as hipóteses do analista.⁸⁴

A inseparabilidade da dimensão interacional da língua, entendida como um recurso de ação social, em relação à dimensão comunicativa da língua, entendida como recurso de troca de informações, revela também as limitações da tipologia de UEs proposta por Chafe. Schegloff e Sacks (1973), por exemplo, analisando os métodos dos participantes para pôr fim a um encontro social, identificam a função interacional de *well* para sinalização do início do processo de encerramento da conversa, revelando uma dimensão da expressão que vai além de seu sentido estritamente referencial (e.g. conectivo de coesão textual). Por outro lado, expressões altamente convencionalizadas como “Bom dia”, que poderiam ser atribuídas ao trabalho interacional de “fazer saudação” em vista do seu alto grau de ritualização, podem também ser interpretadas em seu sentido referencial, como quando, diante de um “Bom dia!” de um colega, respondemos inesperadamente, “Nem tão bom assim”.

Além disso, a atribuição exclusiva de uma função na interação às UEs substantivas e regulatórias, consideradas “bem sucedidas”, em detrimento das UEs fragmentárias (mal sucedidas?), mostra a forte tendência de Chafe em favorecer contextos de fala mais monológicos, em que a negociação de sentidos típica da conversação é minimizada – ainda que Chafe utilize (e afirme favorecer) dados conversacionais. Park (2002), por exemplo, mostra que as UEs truncadas não necessariamente estão relacionadas a problemas de produção, podendo funcionar como recursos convencionalizados em práticas sociais específicas, tal como a busca colaborativa por palavras. Clark (2002), por sua vez, argumenta que as disfluências não devam ser vistas como problemas na fala, mas como soluções metódicas para problemas na fala, tal como a necessidade de coordenação de ações sociais (ver também Clark e Fox Tree, 2002; Clark e Wasow, 1998; Fox Tree e Clark, 1997). Essas dimensões interacionais de toda formulação lingüística, estejam elas correlacionadas a UEs integrais ou fragmentárias, revelam outro fenômeno de que o modelo de comunicação como troca objetiva de informações não pode dar conta.

⁸⁴ Como já mencionado no capítulo anterior, o valor heurístico dessa estratégia deve ser relativizado e não levado às últimas conseqüências. Coulter (1983) mostra que muitas análises no campo da AC são pautadas em estruturas a priori, ainda que levantadas a partir de análises exaustivas de corpora naturais.

Quanto às restrições operantes no processo de produção do discurso, há evidências de que elas de fato existam. Hipóteses sobre limitações de processamento de informação como motivadoras dessas restrições são plausíveis – e, de fato, parecem ajudar a explicar o grande desenvolvimento tecnológico humano (Hutchins, 2005) – mas não excluem a possibilidade de que motivações de ordem interacional também devam ser consideradas (e.g. a maximização de pontos de negociação de sentidos a fim de evitar a quebra da ordem social, cf. nota de rodapé 63 da seção 3.5).

Assim, deslocando-se o olhar mentalista de Chafe para um olhar interacional (i.e. da expressão de estados mentais para o trabalho dos participantes no aqui-agora em face de contingências e propósitos imediatos), a sua proposta ganha nova vida. De fato, as UEs plausivelmente trazem aquilo que, num dado momento da interação, constitui o foco de atenção atual do falante, e que ele espera ser também o foco de atenção atual de seu interlocutor. Isso sugere que a UE desempenhe um papel particularmente importante na construção progressiva de ações sociais na interação, bem como no seu gerenciamento pelos interlocutores. Porém, antes de constituir uma unidade de natureza cognitiva, uma “janela valiosa para o fluxo do pensamento” (Chafe, 1988), as UEs podem ser mais bem entendidas como *formulações* que carregam tanta informação sobre o falante quanto sobre o seu interlocutor, que é a pessoa para a qual aquela formulação foi particularmente ajustada. Uma metáfora mais apropriada para as UEs, então, talvez fosse a de uma “janela valiosa para a *ação social imediata*”.

O campo da AC tem oferecido evidências suficientes de que a tarefa de fazer sentido de uma fala na interação não envolve apenas processamento, codificação e transmissão de informação. Mais do que isso, o interlocutor está engajado numa tarefa de perceber que tipo de *trabalho* o seu interlocutor está levando a cabo ao *selecionar* uma dada instância de fala, em meio a um repertório de opções, num determinado momento da interação. A dimensão informacional dessa fala é mais um instrumento do que uma meta, uma vez que compreender o trabalho interacional de uma fala envolve, entre outras coisas, o estabelecimento de referentes identificáveis e a coordenação da atenção de todos os participantes sobre um mesmo “objeto lingüístico” a cada novo momento. Chafe mostra, nesse sentido, que os recursos prosódicos da língua estão intimamente associados a esses dois importantes processos intersubjetivos.

4.3. O enquadramento da atenção no discurso

Uma proposta sobre as UEs que dá mais alguns passos em direção a uma abordagem interacional é a de Langacker (2001), que recupera a análise de Chafe e a reformula, situando-a num contexto mais amplo em que diferentes dimensões do uso da língua são consideradas. Aproximando-se da metáfora de Chafe de “janelas valiosas para o fluxo do pensamento”, Langacker reformula as UEs como “janelas de atenção” ou “quadros atencionais” (*attentional frames*): gestos atencionais coesos, tipicamente marcados pela entoação, que orientam o interlocutor aos pedaços de informação que devem ser focados a cada novo momento.

Para Langacker, as UEs são estruturas simbólicas assim como os itens lexicais, com um pólo fonológico e um pólo semântico. A diferença reside no fato de que, no pólo fonológico, sua realização se dá principalmente no *nível suprasegmental* – diferentemente dos itens lexicais, realizados principalmente no nível segmental – e no pólo semântico, seu significado se dá principalmente no nível da *estrutura informacional* – diferentemente dos itens lexicais, cujos significados residem principalmente no nível da situação objetiva.⁸⁵ O autor denomina esses níveis *canais de vocalização*⁸⁶ (no que concerne ao pólo fonológico) e *canais de conceitualização* (no que concerne ao pólo semântico).

The semantic pole of an attentional frame belongs to the channel of information structure, and its phonological pole, to intonation. An attentional frame comprises no specific conceptual or segmental content. Its conceptual value resides in the very act of making a single attentional gesture – imposing a single window of attention for the simultaneous viewing of conceptual content. Likewise, its phonological value resides in the very act of imposing an intonational grouping. Despite its abstract nature, an attentional frame is symbolic to the same extent as a lexical item like tree. The

⁸⁵ Apesar do que o termo *situação objetiva* sugere, a gramática cognitiva *não* assume que a realidade possa ser representada na língua de maneira transparente. O que o autor busca apontar aqui é em que nível do pólo semântico (referencial, atencional ou interacional) o canal do conteúdo segmental está principalmente associado, concluindo que se trata do nível referencial.

⁸⁶ O termo “vocalização” de Langacker, referente aos vários canais no pólo fonológico de uma estrutura simbólica, pode parecer infeliz no contexto de análise de uma língua de sinais, por não parecer suficientemente amplo para abranger os gestos manuais e corporais. Apesar disso, Langacker certamente estava ciente do amplo escopo que ele atribuía ao termo, uma vez que o autor incluiu os gestos como um dos vários “canais de vocalização” – em contraste com o uso que foi feito desse mesmo termo na discussão relativa à perspectiva da AC no capítulo 3.

difference is merely that their primary semantic and phonological specifications pertain to different channels (p. 155-6).

Tendo em vista que a prosódia co-ocorre com conteúdos segmentais, Langacker afirma que todo ato de enunciação envolveria um mapeamento entre os domínios atencional (organizado basicamente de forma linear) e estrutural (organizado tanto linear quanto hierarquicamente), através do qual emergem janelas de atenção com parte de uma situação objetiva representada como figura e outra parte, como fundo. Desse modo, diferentes quadros atencionais sobre um mesmo conteúdo conceitual inevitavelmente produzirão enunciados de significado distinto, o que torna a prosódia e a gestualidade tão importantes para a semântica quanto os segmentos da língua.

Langacker argumenta que, assim como ocorre com outras estruturas simbólicas, alguns quadros atencionais podem se tornar arraigados e convencionalizados no sistema lingüístico pela sua recorrência. Recuperando a discussão de Chafe acima, os contornos prosódicos terminais descendentes e ascendentes, bem como os não-terminais, são exemplos de enquadramentos altamente convencionalizados. Em casos mais particulares, alguns desses quadros podem se tornar arraigados no sistema lingüístico junto com as estruturas particulares com as quais eles recorrem. É o caso, por exemplo, da associação entre “contorno não-terminal + contorno terminal”, de um lado, e a estrutura “se p + então q”, de outro (e.g. Se ela disse, então é verdade) (p. 156).

Ao mesmo tempo, Langacker reconhece a necessidade de não se equiparar os quadros atencionais aos constituintes lingüísticos. A existência do padrão “contorno não-terminal + contorno terminal” para a expressão de condicionais do tipo “se p, então q” não impede que essa mesma estrutura co-ocorra com outros enquadramentos atencionais (e.g. um único contorno prosódico terminal para toda a expressão “Se ela disse então é verdade”). Nesse caso, pode-se pensar o quadro “não-terminal + terminal” como um padrão lingüístico *prototípico* para expressão de condicionais, relacionado a outros padrões menos prototípicos disponíveis aos falantes dentro de um repertório de opções.

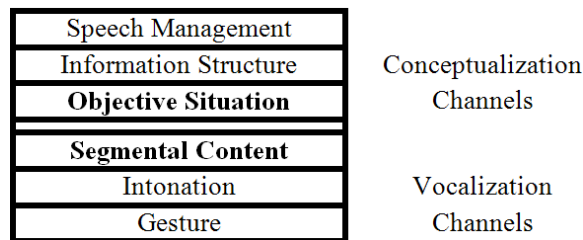


Figura 8. Quadro de Visualização (Langacker, 2001: 146).

A figura 8 acima detalha as diferentes dimensões que, para Langacker, aparecem de maneira menos ou mais explícita na realização de uma UE. Considerando que cada UE corresponda a um “quadro atencional”, Langacker apresenta esquematicamente essas dimensões dentro do que ele chama de “quadro de visualização” (*viewing frame*). Para o autor, num dado evento de uso – realizado por meio de uma UE – estão presentes não apenas os canais de conteúdo conceitual (i.e. a situação objetiva) e da estrutura informacional (e.g. a salientação de certos elementos em relação a outros), mas também o canal do gerenciamento da fala (e.g. questões como a tomada, manutenção e cessão de turnos), cada um desses níveis com uma independência *relativa* em relação aos demais.

Usage events are bipolar, comprising both conceptualizing and vocalizing activity. Moreover (...), at each pole we can reasonably posit multiple channels, which have a *certain amount of independence but are nonetheless coordinated in complex ways*. (...) The various conceptualization and vocalization channels are shown as occurring within the viewing frame. This is, however, *a matter of degree* if we understand the viewing frame as delimiting the general locus of attention (...). At the conceptualization pole, the objective situation is generally more substantive, more concrete, and more at the center of our attention than the other channels. At the vocalization pole, segmental content is comparably salient (p. 146 - ênfase minha)

De acordo com a proposta do autor, então, uma UE pode ser entendida como um *evento de uso* mínimo da língua. O *quadro de visualização* especifica os vários canais de conceitualização e de vocalização envolvidos na realização desse evento. Esse enquadramento deve ser entendido, porém, como uma “questão de grau”, tendo em vista que a separação entre o que emerge como figura no direcionamento da atenção e o que se desloca para o fundo não deve se estabelecer de maneira categórica.⁸⁷

⁸⁷ Em seu trabalho, Chafe já apontava pelo menos três níveis de diferenciação do estatuto mental das informações: ativo, inativo e semi-ativo. Ele reconheceu que essa divisão carecia de maior aprofundamento, mas talvez um ponto crucial da discussão seja o fato da dualidade “objeto focalizado e objeto não-focalizado” delimitar mais um contínuo de elementos capazes de receber diferentes graus de

Ao mesmo tempo, dentro do “lócus geral de atenção” estabelecido por uma nova UE, os diferentes canais também apresentam uma independência apenas relativa, coordenando-se de maneiras complexas. Prototipicamente, a gestualidade apareceria associada ao gerenciamento da fala (troca de turnos e manutenção ou cessão da palavra, etc); a entoação, à estrutura da informação (ênfase, informação dada vs. nova, etc); o conteúdo segmental, à situação objetiva (o conteúdo da fala propriamente dito). Entretanto, isso não impede, por exemplo, que o conteúdo segmental seja utilizado com fins de gerenciamento da fala (e.g. o prolongamento de “é”, para manter a palavra no português brasileiro), ou que a gestualidade seja utilizada para referir-se à situação objetiva (e.g. um gesto de apontamento que ajuda a identificar o referente de um pronome).

Comparando essa proposta com a de Chafe (1994), nota-se que Langacker a elabora de tal modo que uma articulação com a abordagem interacional torna-se mais relevante. Para Chafe, as UEs seriam um esforço dos falantes de tornar as suas conceitualizações disponíveis aos seus interlocutores, permitindo assim reduzir a lacuna, em certa medida irremediável, entre duas mentes distintas. A motivação vislumbrada pelo autor para o processo de produção de UEs seriam as nossas restrições cognitivas de processamento de informação, harmonizadas com a nossa necessidade fisiológica de respiração.

Langacker estende essa idéia e propõe que a motivação das UEs seria a necessidade de *coordenação da atenção dos interlocutores* sobre uma mesma “entidade concebida” a cada novo momento, uma tarefa crucial para o sucesso da comunicação (2001: 144). Essa coordenação da atenção implica a criação de janelas de atenção sobre nossas conceitualizações por meio da imposição de gestos atencionais coesos, simbolizados tipicamente pela entoação, sobre o conteúdo segmental. Esse deslocamento do foco de análise para a organização da atenção na interação nos remete

atenção do que uma escala de pontos discretos. Por exemplo, uma vez enunciada uma nova UE-y, podemos argumentar sem maiores problemas que, digamos, a UE-x imediatamente anterior deslocou-se para o segundo plano da atenção. Contudo, quantas novas UEs serão necessárias até que o conteúdo referente a essa UE-x deixe de ser orientado marginalmente a ponto de não receber mais qualquer atenção? Parece-me que qualquer estipulação nesse sentido seria arbitrária. É uma questão altamente sensível ao contexto imediato, afinal de contas, o quanto a compreensão de uma nova UE dependerá de orientação do interlocutor às UEs previamente enunciadas como ponto(s) de referência(s) (cf. Givón, 1983).

ao papel fundamental da prosódia na coordenação de ações sociais, discutido no capítulo 4.⁸⁸

Ainda assim, o trabalho de Langacker também não escapa da visão da situação de interação como um espaço de transmissão de informações, tal como observamos em Chafe. Para Langacker, a situação objetiva se refere aos conteúdos informacionais que são veiculados de um interlocutor para outro; o enquadramento da atenção e o estabelecimento de relações de figura-e-fundo são formas de organização da língua para o processamento desses conteúdos; e a troca de turnos é uma forma de coordenação social para a veiculação de idéias. Nesse sentido, uma reinterpretação de sua proposta sobre um prisma interacional mais uma vez se mostra relevante, e o que se coloca em debate, novamente, não é a dimensão fonológica do fenômeno analisado, mas a semântica.

A visão a ser reformulada é, como no caso de Chafe, a de que a interação é essencialmente performática: um domínio de ação social. Cada um dos níveis descritos por Langacker é empregado para viabilizar não apenas a troca de informações, mas o estabelecimento de relações sociais, o que torna o quadro de visualização todo impregnado de significação social. A prosódia, por exemplo, não é apenas um recurso que permite ao receptor processar uma quantidade “digerível” de informações, como afirma o autor (Langacker, 2001: 154). Ela é um dos recursos de formulação que *revela* aos participantes, entre outras coisas, o escopo de interpretação da fala imediata, decisão que se mostra altamente conseqüente para a interpretação sobre o tipo de ação que *está* sendo realizada e as que *não* estão sendo realizadas (e.g. exemplo de “I gave up

⁸⁸ Tanto a análise de Langacker quanto a de Chafe deixam em aberto a questão de por que essas “janelas de atenção” devam ser tipicamente marcadas por meio da entoação – e não, por exemplo, por meio de agrupamentos de fala sem qualquer modulação, separados por cesuras. Uma possível resposta a esse tipo de questionamento aparece na tese que Dwight Bolinger perseguiu ao longo de sua carreira: a de que a entoação integre um complexo gestual que possui uma motivação primitiva de expressão da emoção. Para o autor, a entoação “dá suporte à gramática – em alguns casos pode ser indispensável a ela – mas não é, em última instância, gramatical” (Bolinger, 1985: 106). O fato de que as crianças são capazes de produzir determinadas modulações de pitch antes mesmo de poderem compreendê-las na fala dos adultos – um caminho inverso ao da aquisição, que procede da compreensão para a produção – seria uma evidência de que as variações de pitch para cima ou para baixo estão mais correlacionadas ao nível de interesse e entusiasmo do que a uma intenção lingüística (Bolinger, 1983: 160).

Para Bolinger (1983), então, o pitch sobe mais ou menos de maneira gradiente e icônica de acordo com o que nos empolga, e o interlocutor interpreta tais variações como expressão de uma informação “importante”, “em foco”, “surpreendente”, “nova” (p. 160). Contornos entoacionais ascendentes e descendentes seriam igualmente gradientes e icônicos – ainda que *metaforicamente* icônicos – assinalando cesuras que indicam, respectivamente, níveis de abertura e fechamento, continuação e finalização, dependendo do grau com que sobem ou caem (Bolinger, 1998: 48).

smoking cigarettes” na nota de rodapé 74). E o propósito das diferentes possibilidades de estruturação da fala, a partir da exploração de recursos em vários níveis de análise, é o de permitir a adaptação dos indivíduos às contingências das situações de interação em face de seus propósitos práticos imediatos (Edwards, 1998).

Essa discussão mostra o quão relevantes mutuamente são, de um lado, o conceito de “formulação” da etnometodologia e da AC (Garfinkel e Sacks, 1970), e de outro, o conceito de “estruturação ideacional” (*construal*) da gramática cognitiva (Langacker, 1987). Na visão da gramática cognitiva, cada forma particular de estruturar uma dada fala traz consigo uma forma igualmente particular de conceber uma dada situação, não havendo duas instâncias fonologicamente distintas de fala que apresentem um mesmo conteúdo conceitual.⁸⁹ Essa é, em essência, a proposta por detrás do conceito de formulação, segundo o qual qualquer prática de glosagem sobre uma dada instância de fala (i.e. uma paráfrase) apenas acrescenta novas camadas de significação à fala glosada, ao invés de avançar no sentido de “explicar” ou “esgotar” o seu significado (cf. nota de rodapé 65 da seção 3.5). Essa análise revela o valor moralmente conseqüencial de qualquer que seja a escolha de formulação dos participantes, de tal forma que tanto a necessidade quanto a dispensabilidade da prática de glosagem estarão submetidas aos requerimentos da situação, de acordo com os propósitos práticos dos participantes.

Mais do que similares, no entanto, entendo que esses conceitos sejam complementares, o que torna a sua articulação frutífera para ambas as áreas de investigação. O ponto forte do conceito de estruturação ideacional está na clareza e no detalhamento das diferentes dimensões que constituem as expressões lingüísticas,⁹⁰ bem

⁸⁹ Um exemplo ilustrativo que Langacker (1991) discute é o de frases do tipo: “O João se parece com a Maria” e “A Maria se parece com o João”. Embora possa parecer que ambas se constituam em paráfrases perfeitas, apresentando o “mesmo conteúdo” com formas distintas, o autor argumenta sobre a necessidade de reconhecer, no mínimo, uma diferença de perspectiva e proeminência entre cada uma das frases: no primeiro caso, a Maria é a base de comparação e a fala é sobre o João; ao passo que, no segundo caso, o João é a base de comparação e a fala é sobre a Maria. A ilusão da paráfrase perfeita se desmancha quando deixamos de olhar para as frases de maneira descontextualizada e passamos a investigá-las em situações concretas de uso, onde a troca de uma por outra simplesmente não é possível.

⁹⁰ A estruturação imposta por uma dada expressão lingüística envolve outras dimensões além da proeminência: i) nível de especificidade/esquematicidade (e.g. referir-se a um *canino*, e não a um *cachorro*, ou a um *pastor alemão*, etc); ii) o nível de escopo da predicação (e.g. em **O braço tem cinco dedos*, “cinco dedos” viola o escopo imediato de predicação designado pelo sujeito “braço”); e iii) a perspectiva (e.g. **Eu vim lá* e **Eu fui aqui* são infelizes porque a direcionalidade intrínseca aos verbos é incompatível com a direcionalidade dos pronomes (para um resumo da proposta, ver Langacker, 1986).

A meu ver, o conceito de estruturação ideacional de Langacker poderia ser ainda mais enriquecido se

como na caracterização formal daquilo que, no discurso espontâneo, constitui uma unidade mínima de estruturação: as UEs. O ponto forte do conceito de formulação, por sua vez, está em apontar a motivação fundamentalmente social da multiplicidade de opções de estruturação da fala, bem como em explicar como é que, diante de tantas opções, os participantes são capazes de compreender o sentido das falas uns dos outros. A articulação dessas propostas oferece, portanto, uma fundamentação fonológica e semântica sólida o estudo das unidades mínimas do discurso.

4.4. O papel da salientação na experiência humana

A organização de figura-e-fundo parece ser um dos princípios mais fundamentais da experiência humana. Um exemplo de manifestação típico dessa organização está na percepção – domínio no qual esse fenômeno foi mais extensivamente estudado por psicólogos e neurocientistas. Grosso modo, como aponta Langacker (2000), a visão se organiza de tal maneira que um observador estrutura, diante de si: a) um foco central de atenção visual (i.e. a figura); b) uma área periférica na qual esse foco se situa e sobre a qual a acuidade visual é medial (i.e. a parte mais saliente do fundo); e c) um campo máximo de visão, onde a acuidade visual é mais limitada (i.e. a parte menos saliente do fundo).

Essa organização não se restringe à visão, mas se estende à percepção de modo geral. Por exemplo, para conversarmos com um ou mais amigos numa festa com música e muitas outras pessoas falando, precisamos ser capazes de direcionar o nosso foco de atenção para a voz de quem fala, mantendo as vozes dos demais participantes da conversa no fundo de nossa atenção, e deixando de atender diretamente a todos os demais ruídos da casa; caso contrário, a ocorrência simultânea de diferentes sons e ruídos tornaria a conversação em tais contextos inviável.

Os exemplos acima sugerem duas características importantes da organização de figura-e-fundo em nossa experiência cotidiana (Taylor, 2002: 10). Em primeiro lugar, essa relação pode ser altamente dinâmica. A visão, por exemplo, tende a fixar-se muito pouco tempo num único ponto, deslocando-se continuamente e, portanto, tornando o que agora é figura (i.e. o foco de nossa atenção) em fundo num momento subsequente. Além disso, vários níveis de organização de figura-e-fundo estão presentes

fosse dada uma maior consideração à dimensão prosódica e gestual da fala, na linha de análise, por exemplo, de Bolinger (1983).

simultaneamente, e qual nível se tornará mais saliente depende da relação que estabelecemos com o objeto em questão. Por exemplo, ao lermos um texto, as letras pretas emergem como figura sobre uma página branca de fundo, mas se deslocarmos nosso nível de observação para a própria folha de papel, ela poderá emergir como figura em relação a um fundo caracterizado, por exemplo, pela mesa que lhe serve de suporte.

No que se refere à língua, os trabalhos de Talmy (e.g. 1983; 1996) e Langacker (e.g. 1987), em especial, demonstram que a relação de figura-e-fundo se manifesta em diferentes níveis de análise, constituindo-se num elemento central de sua estruturação. A presente seção irá discutir, então, de que modo essa organização se manifesta em alguns desses níveis – as UEs sendo uma das manifestações do alinhamento figura-e-fundo no nível do discurso – e irá sugerir que vários fatos descritos nos estudos da interação possam também ser interpretados, pelo menos em parte, como motivados por essa forma básica de organização da experiência.

A emergência de um elemento como proeminente em relação a outro(s) elemento(s) que se desloca(m) para um segundo plano constitui o processo conhecido como salientação (do inglês *foregrounding*). Como discutido na seção anterior, uma das premissas da gramática cognitiva é a de que toda expressão lingüística traga consigo uma *estruturação ideacional* específica, isto é, cada expressão particular envolva a imposição de uma concepção imagética subjetiva (*subjective imagery*) sobre uma dada situação objetiva. Dentre as diferentes dimensões que constituem essa estruturação, então, a proeminência relativa de seus elementos está entre as mais fundamentais (Langacker, 2001: 158).

Apesar disso, figura e fundo são conceitos tão gerais que a sua relevância para o entendimento dos diferentes âmbitos da experiência humana depende de uma especificação dos níveis de análise em que tal organização se manifesta. Uma primeira manifestação do fenômeno da salientação na língua, então, se revela nos conceitos de *realce*, *base* e *domínio* (Langacker, 1987). Tais conceitos podem ser ilustrados considerando-se o significado da palavra “hipotenusa”, que designa o lado oposto ao ângulo reto de um triângulo retângulo. Embora o elemento *realçado* (designado) pela palavra seja um dos lados do triângulo (i.e. o lado oposto ao do ângulo reto), não é possível entendermos plenamente a expressão “hipotenusa” sem situarmos esse elemento designado em relação à sua *base*: a concepção de um triângulo retângulo. O realce de uma expressão é, portanto, um elemento da base que emerge como saliente em

relação aos demais elementos. Entender o que é um triângulo retângulo exige ainda noções mais gerais de geometria, tais como a definição de triângulo em relação a outras figuras geométricas, a noção de plano, entre outras. Esse conhecimento mais amplo ligado a uma expressão é o que Langacker chama de *domínio*.

Os conceitos de realce, base e domínio se aplicam não somente a palavras, mas a qualquer expressão lingüística. Tomemos, como um segundo exemplo, o sintagma: “O homem de calça azul ao lado da porta”. Nele, o realce é “o homem”. A evidência disso é que qualquer coisa que predicarmos sobre a expressão como um todo (e.g. “O homem de calça azul ao lado da porta *é feio*”) será, em última instância, uma predicação sobre o seu realce, no caso “o homem” (Taylor, 2002: 193). É esse elemento, portanto, que está sendo designado por essa expressão e que acaba salientado, orientando o nosso foco de atenção principal.

No entanto, não é possível entender esse sintagma nominal complexo plenamente sem entender os sintagmas preposicionais (“de calça azul” e “ao lado da porta”) que o integram. Como discutido no exemplo da hipotenusa, o realce é um elemento da base que se destaca, e nesse segundo exemplo a base envolve também as informações relativas a esses sintagmas preposicionais, que receberão um foco de atenção secundário. Por fim, há que considerar ainda o conjunto de conhecimentos que serve de suporte para o entendimento tanto do realce quanto da base: o domínio. No caso, o conhecimento sobre habitação, vestuário e cores são alguns dos elementos envolvidos na conceitualização da expressão que se fazem necessários para a compreensão desse sintagma complexo. Tais elementos terão um foco de atenção ainda mais marginal do que a base, o que *não* torna o domínio um conceito menos importante.⁹¹

Dando um segundo passo na análise das manifestações lingüísticas da salientação, então, Langacker (1987) analisa a estrutura do realce em termos dos conceitos de *trajetor* e *objeto de referência*, que ajudam a entender a organização de expressões em torno de itens lexicais de natureza relacional. Por exemplo, na

⁹¹ Por exemplo, no decorrer da interação, elementos que foram projetados pelo falante corrente para receberem um foco de atenção secundário ou ainda mais marginal podem, no processo natural de negociação entre os interlocutores, ser deslocados para o foco de atenção num momento subsequente. O fato de a dimensão da “base” de um conceito ser em alguns trabalhos referida como a dimensão do “domínio imediato” mostra que esses conceitos não devem ser assumidos como categóricos.

conceitualização da relação expressa pela preposição “sobre”, duas entidades estão esquematicamente envolvidas: uma está por cima da outra. O trajetor (TR) e objeto de referência (OR) são os termos dados por Langacker a essas entidades esquemáticas do realce, o OR representando o ponto de referência do realce (com foco de atenção secundário) e o TR representando o objeto que se localiza em relação a esse ponto de referência (com foco de atenção principal).⁹² Para a preposição “sobre” ser integrada num sintagma como “o livro sobre a cadeira”, por exemplo, a preposição teve as suas duas entidades esquemáticas (TR e OR) elaboradas por instâncias específicas: “a cadeira” elabora o OR, formando o constituinte “sobre a cadeira” e “o livro” elabora o TR dessa expressão composta, formando o sintagma completo.⁹³

Dessa maneira, os conceitos de realce, base e domínio, de um lado, e trajetor e objeto de referência primário e secundário, de outro, se revelam como manifestações do processo de salientação no nível do enunciado e dos constituintes mais básicos que o compõem. Ampliando mais o escopo de análise, então, percebemos que os próprios enunciados estabelecem entre si diferentes relações de proeminência no âmbito do discurso. Essa outra dimensão do fenômeno de salientação se manifesta no conceito de UEs, discutido nas seções anteriores.

Clark (2004) articula explicitamente essa visão sobre as UEs em sua análise sobre a relevância da temporalidade na interação. O autor chama de “pacotes” (*parcels*) os pedaços de fala delimitados por uma única UE, afirmando que a alocação temporal desses pacotes é uma dimensão crucial do significado das expressões lingüísticas na fala espontânea:

When I place a book on a table, a pan in an oven, or a coat on a hook, I place one object *with respect to* another – a figure with respect to a ground... What speakers mean by the placement of a parcel depends on what is figure and what is ground. When an official at a track meet starts a ten-thousand kilometer race by saying, ‘Ready... set... go!’, he uses the realization of ‘go’ to indicate the beginning of the race. To do that, he places ‘go’ (the figure) with respect to the ‘ready ... set’ (the ground)... If [he] had placed the

⁹² O termo “trajetor” foi formulado em referência a uma situação típica de trajetória, na qual uma “entidade percorre um trajeto” (*trajector*) e se movimenta em relação a um ou mais pontos fixos de referência (*landmark*).

⁹³ Esse exemplo, baseado em Taylor (2002), não discute o papel dos artigos na conceitualização da expressão.

same phrase before ... he might have been taken to mean something different (p. 1-2, *itálicos no original*).

Ainda segundo Clark, os falantes reconhecem que a maioria das UEs na fala espontânea possui um padrão de alocação não-marcado. Esse padrão prediz que o início de uma nova unidade deva ocorrer ao término da unidade anterior, ainda que a projeção desse término esteja submetida à indeterminação e contingencialidade da fala espontânea – como discutido no capítulo 3. Desse modo, os falantes optam pela alocação padrão a menos que eles queiram significar alguma coisa em especial por meio de uma alocação alternativa.

O trabalho de Talmy (1996) mostra que um dado conteúdo segmental que co-ocorre com uma UE pode estabelecer relações de figura-e-fundo não apenas em relação às unidades prévias do próprio discurso, mas também em relação a partes de conceitualizações complexas (*frames*) que certas expressões evocam. Por exemplo, o frame de um evento caracterizado por uma trajetória prototípica envolve, do ponto de vista conceitual, pelo menos três etapas distintas: o início da trajetória, seu percurso e o seu ponto final. Na estruturação de uma situação objetiva por meio da codificação lingüística, então, o falante pode optar por salientar uma ou mais partes dessa trajetória de acordo com os requerimentos da situação, os seus propósitos particulares e, possivelmente, sob restrição de limitações na produção.

Para ilustrar esse processo, Talmy apresenta um exemplo do seguinte tipo: “Uma das caixas da carga da aeronave caiu [para fora do avião]₁ [pelo ar]₂ [no oceano]₃”. As expressões lingüísticas demarcadas pelos colchetes 1, 2 e 3 indicam, respectivamente, o ponto inicial, o percurso e o ponto final da trajetória da caixa. Os colchetes indicam ainda que cada uma das expressões neles contidas pode ser explicitamente mencionada ou não, opção que irá trazer para o foco da atenção, ou deslocar para um plano atencional secundário, certos aspectos de trajetória do objeto que são evocados como um todo pela expressão de um verbo como “cair”. Por exemplo, em “Uma das caixas da carga da aeronave caiu no oceano”, o falante salienta o ponto final ao mesmo tempo em que omite o ponto inicial e medial da trajetória, que permanecem implícitos e subespecificados na periferia da atenção. Talmy argumenta que o mesmo processo se manifesta com vários tipos de frames de eventos que apresentam uma estrutura cognitiva complexa e estável.

Em todos os trabalhos acima discutidos, destaca-se o modo como as alternativas lingüísticas disponíveis aos falantes para a formulação de uma situação contribuem para o gerenciamento da atenção no discurso, com o falante empregando diferentes estratégias formais a fim de direcionar a atenção de seu interlocutor para um aspecto específico de uma conceitualização e não outro. Esse trabalho de gerenciamento da atenção, que opera inclusive nos níveis mais micro-analíticos da gramática (e.g. as relações entre trajetor e objeto de referência, que organizam a composição das expressões mínimas da língua), pode, com o deslocamento do foco de investigação para níveis mais macro-analíticos, se revelar igualmente pertinente à organização de outras dimensões da experiência humana, como é o caso da interação.

De fato, observando os trabalhos voltados para a interação, salta aos olhos o modo como vários conceitos centrais parecem estar intimamente relacionados com o fenômeno da salientação. Por exemplo, Goffman (1964) define o *encontro* dentro de uma situação social como um “agrupamento ecológico no qual os participantes *orientam-se uns em relação aos outros* ao mesmo tempo em que *se apartam dos que estão presentes na situação mas não oficialmente no encontro*” (p. 135, itálicos meus). O encontro seria, portanto, a circunstância em que “duas ou mais pessoas ... cooperativamente se ratificam umas às outras como mantenedores autorizados de um único, ainda que móvel, *foco de atenção cognitivo e visual*” (p. 135, itálicos meus).

Goffman (1961) mostra também que, dentro dessa atenção já seletiva que as pessoas dedicam umas às outras em uma situação social, há um segundo plano de seletividade, agora num nível mais macro-social. Ocorre que, devido às inúmeras identidades potencialmente relevantes num dado encontro, os falantes possuem formas de sinalizar qual *aspecto identitário está sendo salientado num dado momento* e quais devem permanecer à margem (Erickson, 2004: 148-9). Essa salientação de certos aspectos identitários é uma dimensão inextricável das formulações, que estabelecem, de acordo com os propósitos práticos imediatos, o próprio falante e/ou seus interlocutores como pertencentes a determinadas categorias sociais (Sacks, 1972, 1992a: 41-48).

Diferentemente de Goffman, os estudos da AC herdaram de sua base etnometodológica uma resistência teórica particular em tratar os fenômenos da interação em termos “cognitivos” (e.g. J. Coulter, 1991; Lynch, 2006; Molder e Potter, 2005; Schegloff, 1991b); apesar disso, alguns conceitos levantados nos estudos da AC a meu ver reforçam a relevância do fenômeno da salientação para a análise da interação –

independentemente da natureza “cognitiva” ou “social” que possa e/ou deva ser atribuída a tal fenômeno.

Dentro do escopo do “encontro” social de que Goffman fala, por exemplo, em que “duas ou mais pessoas” ratificam-se como mantenedoras de um foco cognitivo e visual, SSJ identificam o procedimento sistemático de alocação de um *único falante a cada vez* (p. 19), que opera independentemente do número de participantes da conversação. É plausível especular que esse princípio seja, pelo menos em certa medida, resultado do fenômeno da salientação, tendo em vista que a pressão social na interação por contribuições seqüencialmente implicativas (i.e. que respondam às anteriores de maneira temporalmente coordenada e interacionalmente relevante) restringe fortemente as opções de ação e direcionamento da atenção dos interlocutores diante de um turno em curso – na medida em que os participantes não queiram ser moralmente responsabilizados por “não falar nada com nada”, “não estar nem aí”, “ficar tirando sarro”, entre outras possíveis formulações freqüentemente dirigidas a contribuições consideradas inadequadas à interação.

Uma outra evidência para essa relação entre salientação e interação encontra-se no fenômeno conhecido como *cisma*, já assinalado por SSJ e retomado em maior profundidade no trabalho de Egbert (1997). De acordo com esse fenômeno, conversações com dois ou três participantes se desenvolvem de maneira bastante similar, isto é, por meio da alocação dos turnos a uma única parte a cada momento, o que resulta em configurações sempre binárias de participação – e, no caso de conversas com três participantes, em sobreposições que revelam a competição de mais de um participante pelo estatuto restrito de interlocutor (Schegloff, 2000: 8-10). Porém, em conversações com quatro ou mais participantes, em que há interlocutores em potencial para se orientarem e responderem a mais de um falante simultaneamente, observa-se a prática comum do *cisma*, quando uma única conversação se desdobra em duas paralelas, oferecendo assim um alívio para o conflito entre a quantidade de interlocutores potenciais e o caráter binário de alocação das participações. Desse modo, quanto maior o número de participantes numa interação – partindo do número mínimo de quatro – maior a probabilidade de ocorrência de cismas.

Ainda que breve, a discussão acima serve para ilustrar a importância do conceito de salientação – que envolve a imposição de um elemento como figura em relação a outro(s) elemento(s) como fundo – para a organização de domínios diversos da

experiência humana, tais como a percepção, a cognição, a língua e a interação. Observando-se em especial o que ocorre no âmbito da interação, a discussão sugere que, mais do que um mero *recurso* de organização da experiência, a relação entre figura e fundo se imponha como um possível imperativo em tarefas que demandam um grande investimento atencional e responsivo, como é o caso da conversação.⁹⁴ A prevalência do fenômeno de salientação, associada a sua relação estreita com a coordenação da atenção, colocam-se, portanto, como argumentos adicionais para uma consideração especial da prosódia nas análises sobre a projetabilidade e a coordenação das ações sociais na conversação.

⁹⁴ Como visto na discussão acima, Chafe coloca uma grande ênfase na hipótese de que a aparente limitação das UEs reflita restrições de processamento cognitivo. Considerando-se as diferentes manifestações do um-de-cada-vez discutidas nesta seção, uma reflexão mais apressada poderia ser a de que tais manifestações sejam novas evidências dessa restrição. Contudo, a questão das possíveis restrições relativas aos processos atencionais, debatida no âmbito da psicologia, está longe de alcançar uma resposta mais definitiva.

Embora a analogia sobre o funcionamento do olhar – em termos de uma região focal, uma periférica e uma que foge do alcance perceptual – sirva como uma ilustração geral pertinente para a reflexão sobre o modo de operação da atenção nos humanos, várias questões fundamentais permanecem: por que é necessário atender a apenas um objeto especificamente? O que acontece com os objetos não atendidos? Porque certos objetos e ações complexas exigem atenção e não outros? (Underwood, 1993).

A partir dos trabalhos pioneiros de William James, ainda no século XIX, e de alguns pesquisadores até a metade do século XX, a primeira resposta a essas questões foi similar à de Chafe: a de que devesse haver uma espécie de filtro atencional, de modo que, uma vez atendido um objeto, outros fossem automaticamente excluídos do processamento. Estudos revelando claros efeitos de informações não-atendidas sobre as informações atendidas, porém, exigiram reformulações desse modelo nas décadas subsequentes. Uma das reformulações foi a de que as informações não-atendidas não seriam excluídas do processamento, mas simplesmente atenuadas.

O aprofundamento dessa discussão fez emergir no campo, então, uma problematização do lócus atencional (seja de um possível filtro, ou de uma possível atenuação de informação) ao longo dos diferentes estágios de processamento. Como mostra Underwood (1993), vários estudos das décadas de 70 e 80 argumentaram que o problema central se revelava em estágios mais avançados do processo, como a seleção de uma *resposta* a um estímulo, e não na sua identificação e análise perceptual. Outros autores procuraram destacar que, na busca desse lócus, não apenas os processos de baixo-para-cima deveriam ser considerados, mas também os de cima-para-baixo, como o papel de expectativas provenientes de experiências físicas e culturais prévias. Por fim, alguns autores têm chamado a atenção para a necessidade de considerar o processo de automatização de certas tarefas cognitivas, tendo em vista que a prática contínua de uma mesma tarefa parece, com o decorrer do tempo, eximir os indivíduos de recursos de processamento para a sua execução (e.g. o processo de aprender a guiar um carro).

Assim, embora haja evidências robustas de que as pessoas apresentem limitações para a realização de mais de uma tarefa cognitiva simultaneamente (ver síntese crítica em Pashler, 1994), a hipótese de restrições cognitivas por detrás dos processos atencionais se mostra altamente complexa e, na verdade, tangencial ao argumento central deste estudo. Mais diretamente relevantes aqui são as evidências, a serem discutidas na seção 4.5 a seguir, de que a determinação de um foco atencional conjunto desempenhe um papel fundamental na evolução de várias espécies, em especial a humana, constituindo-se num dos pré-requisitos mais importantes para a aquisição das línguas naturais.

4.5. Salientação, gestualidade e atenção conjunta

Embora seja possível argumentar que a prosódia desempenhe um papel central para o estabelecimento da atenção conjunta entre os humanos, a exploração do trato vocal é apenas uma das formas – embora possivelmente a forma mais importante no caso de nossa espécie – a contribuir para a realização da interação espontânea face-a-face. Outros recursos que também integram os processos atencionais, e que são inclusive anteriores à língua tanto ontogenética quanto filogeneticamente, são as várias formas de gestualidade, como os gestos de apontamento, o olhar, a orientação da cabeça e do tronco.

De fato, no âmbito da psicologia e da antropologia evolutiva, estudos centrados na análise de como a atenção social é estabelecida entre primatas, incluindo humanos, têm se concentrado principalmente nessas dimensões mais primitivas da gestualidade (para duas sínteses críticas, ver e.g. Emery, 2000 e Langton et al., 2000). Tratando mais especificamente do olhar, Langton et al. (2000) destacam o papel que esses recursos gestuais desempenham na evolução de determinadas espécies:

Humans and many other species tend to look at things in their environment that are of immediate interest to them. You might be the recipient of another's gaze, for instance, because you are a potential meal, a mate or simply because you are someone with whom they would like to interact. Individuals who are able to detect rapidly when they are the object of another's attention, and who can analyse exactly where another's gaze is directed therefore have considerable adaptive advantage (p. 51-2).

Em humanos, pesquisas têm demonstrado que o redirecionamento do olhar para uma região específica parece desencadear mudanças reflexivas (obrigatórias) de orientação por parte de seu(s) observador(es), levando-o(s) a deslocar a sua atenção automaticamente para essa mesma região (e.g. Langton e Bruce, 1999). O efeito desse deslocamento é a facilitação do processamento de qualquer objeto-alvo que apareça nesse local. Nesse sentido, Langton et al. (2000) destacam que as pistas gestuais, tal como o direcionamento do olhar, se diferenciam de pistas provenientes de fenômenos naturais, tal como o estrondo de um trovão: ao passo que as primeiras “empurram” a atenção do interagente para a região indicada, as últimas “puxam” a atenção dos indivíduos para as regiões onde elas se manifestam (p. 55). Essa distinção revela o caráter intimamente social, intersubjetivo, da gestualidade humana.

A importância do olhar nos processos atencionais, porém, não deve ser desvinculada de outras formas de gestualidade, como a direção da cabeça e do tronco. Alguns autores, como Perrett e Emery (1994), privilegiam o olhar sobre outras pistas gestuais, argumentando que, quando as informações de diferentes articuladores entram em conflito (e.g. a cabeça se volta para o lado e os olhos para frente), uma rede de conexões inibitórias no indivíduo observador faria com que a informação do olhar sobrepujasse a da cabeça, e ambas a do tronco. Desse modo, uma hierarquia que parte da orientação do olhar, passando pela da cabeça, até a postura do tronco, poderia ser postulada.

Langton et al. (2000), contudo, argumentam a favor da necessidade de considerar algum nível de interação entre esses diferentes articuladores, apontando que tronco e cabeça, quando em conflito com o olhar, devem contribuir em alguma medida para o cômputo da direção da atenção. Por exemplo, na figura 9 abaixo – originalmente ilustrada por William Wollaston, em 1824, e reproduzida em Langton et al. (2000: 55), o olhar parece estar direcionado à direita do leitor na imagem (a), e na direção do leitor na imagem (b); contudo, ao tampar com as mãos a região abaixo dos olhos, nota-se que o olhar, em si, apresenta exatamente a mesma orientação em ambas as imagens.

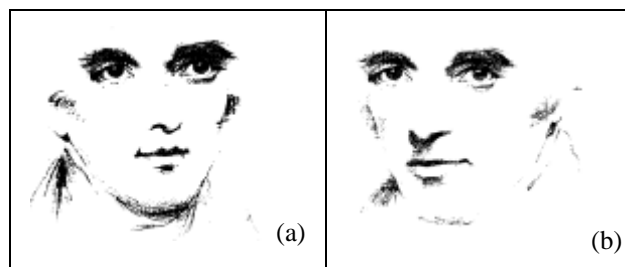


Figura 9. Influência da orientação da cabeça sobre a direção do olhar ilustrada por Wollaston.

O processo de direcionamento atencional através da orientação de diferentes articuladores corporais envolve uma série de etapas distintas que aparece ilustrada na figura 10 abaixo – adaptada de Emery (2000: 590). A primeira etapa importante se refere à distinção entre a *observação mútua* (quando dois indivíduos A e B revelam orientação mútua) e o *olhar divergente* (quando o indivíduo A se orienta a B, que no entanto se orienta a algum outro lugar). Em segundo lugar vem o *acompanhamento do olhar*, quando o indivíduo A percebe a orientação distinta de B e desloca seu olhar para o ponto em que B está olhando. A etapa seguinte, da *atenção conjunta*, se diferencia da de acompanhamento do olhar na medida em que tanto A quanto B atendem ao mesmo

objeto, e não meramente ao mesmo local. Já a etapa de *atenção compartilhada* é um pouco mais complexa que a anterior, pois envolve, simultaneamente, a atenção de A e B a um mesmo objeto e em relação a si próprios – havendo a necessidade de alguma forma de checagem mútua de que ambos estão atendendo ao mesmo objeto. Por fim, a *atribuição de estados mentais* envolve a utilização, por parte de ambos os indivíduos, de sua empatia e experiências prévias a fim de determinar que, se o outro está atendendo a um dado objeto, ele assim o faz porque deseja realizar algo com esse objeto ou porque traz alguma crença sobre esse objeto.

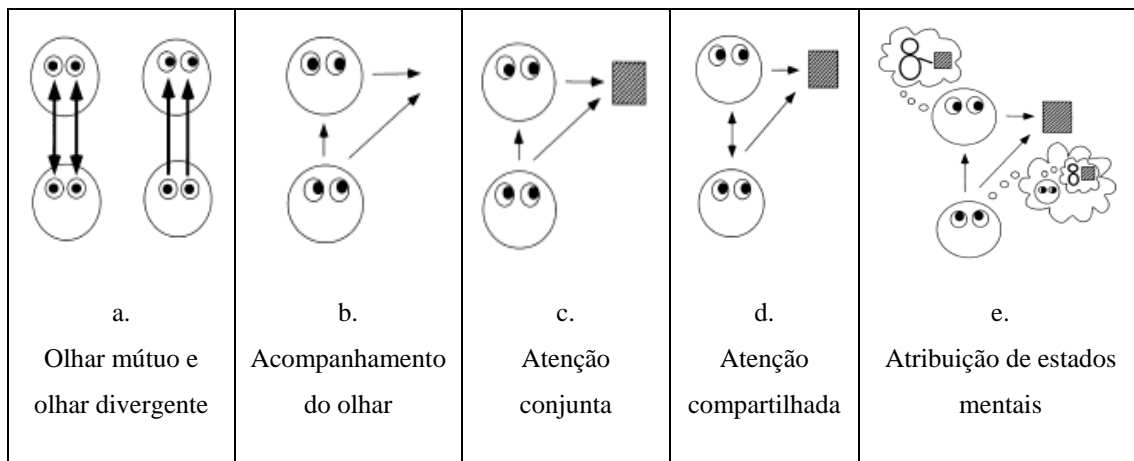


Figura 10. Etapas no processo de direcionamento atencional descritas por Emery (2000).

Há um contínuo debate sobre em que medida outros primatas seriam ou não capazes de atingir as etapas mais elaboradas de coordenação atencional, como a atribuição de estados mentais. Além disso, estudos desse processo com seres humanos têm revelado que a própria criança passa gradativamente por essas diferentes etapas ao longo de seu desenvolvimento. Os trabalhos de Michael Tomasello e colegas, em especial, têm destacado a função primordial dos processos de atenção conjunta⁹⁵ para o desenvolvimento cultural humano e, com relevância mais direta para a presente pesquisa, para a aquisição da língua (e.g. Tomasello e Farrar, 1986; Tomasello e Haberl, 2003; Tomasello et. al., 1993; e para sínteses das propostas, ver Tomasello, 1992b, 2000a, 2000b).

⁹⁵ O termo “atenção conjunta” costuma ser utilizado na literatura para fazer referência ao processo de coordenação da atenção de maneira geral. O presente trabalho segue essa tendência e também fará uso desse termo em referência ao processo como um todo, reservando referências explícitas a uma ou outra etapa específica apenas quando a sua discussão se mostrar particularmente relevante.

Tais trabalhos têm demonstrado que, inicialmente, até por volta de nove meses de idade, a criança possui muito poucos recursos para estabelecer uma atenção conjunta com os seus interagentes, o que torna o papel do adulto mais determinante no estabelecimento da interação. Tomasello e Farrar (1986) mostram, por exemplo, que a aquisição de linguagem nesse período está mais positivamente relacionada aos momentos em que o adulto segue a orientação espontânea da própria criança do que aos momentos em que o adulto tenta redirecionar a atenção da criança para um objeto não-atendido por ela naquele instante.

Entre nove e doze meses de idade, porém, o comportamento da criança começa a se modificar. Uma gama de habilidades atencionais começa a despertar nesse período, tais como o acompanhamento do olhar e dos gestos de apontamento de outros indivíduos, a imitação da ação deles sobre os objetos, e tentativas de manipular a sua atenção apontando e segurando objetos para eles (Tomasello, 2000a: 64). Aqui, novamente a correlação entre atenção conjunta e aquisição da língua se mostra fundamental, na medida em que crianças que desenvolvem habilidades atencionais mais cedo também apresentam uma aquisição de habilidades lingüísticas mais precoce (Carpenter et al., 1998).

Com o passar do tempo, o desenvolvimento das habilidades atencionais na criança vai criando condições cada vez maiores para que ela se torne mais ativa na tarefa de determinar o foco de atenção dos adultos. Por volta dos dois anos de idade, as crianças são capazes de seguir “ordens” para deslocar seu foco de atenção atual e também de focalizar o mesmo objeto que os adultos estão atendendo, podendo aprender palavras relativas tanto a esses objetos quanto às “ordens” a eles relacionadas (Tomasello, 2000a: 71).

Mais complexa, porém, é a tarefa de aprendizagem de verbos e palavras referentes a ações. Tais palavras são empregadas na maioria das vezes antes que as ações por elas referidas sejam realizadas, o que impede o estabelecimento de uma atenção conjunta sobre a ação *no preciso momento* em que ela é nomeada. Apesar disso, evidências naturalísticas e experimentais mostram que é em contextos de antecipação (i.e. quando a palavra é nomeada antes de a ação ocorrer, como em “Chuta a bola pro papai”) que a criança mais bem aprende essas palavras, e não nas situações em que a ação referida e a palavra co-ocorrem temporalmente (Tomasello e Kruger, 1992). Para esses autores, esse fato revela o papel fundamental dos ambientes e atividades culturais

estruturadas, que, quando recorrentes na experiência da criança, permitem a ela antecipar ações subseqüentes com grande sucesso. Quando a criança torna-se capaz de fazer esse tipo de antecipação, então, ela já revela estar operando num nível mais elaborado, de *atenção compartilhada*, e não simplesmente de *atenção conjunta* com seus interlocutores.

Em última instância, a proposta desses autores é a de que o desenvolvimento da língua pela criança depende do entendimento de que os sons produzidos pelos humanos em contextos interacionais servem para manipular a sua atenção ou comportamento para alguma terceira entidade relevante – a *intenção comunicativa* (Tomasello, 1998) –, seja essa entidade do mundo físico e cultural imediato ou do mundo conversacional.⁹⁶ No que diz respeito ao uso lingüístico estrito senso nessas interações adulto-criança, Tomasello segue a proposta geral de teóricos funcionalistas e argumenta que a unidade psicolingüística fundamental nesse processo seria o *enunciado*: “um ato lingüístico no qual uma pessoa expressa em relação a outra, *sob um único contorno entoacional*, uma intenção comunicativa relativamente coerente num contexto comunicativo” (2000a: 63, ênfase minha).

Nesse ponto, então, a discussão sobre o estabelecimento da atenção conjunta nos permite reestabelecer mais explicitamente o elo com a discussão anterior deste capítulo, a respeito da função primordial de salientação da prosódia e do seu papel no gerenciamento da atenção na conversação. A prosódia pode ser entendida como um dos recursos, dentre outros de natureza gestual, por meio dos quais os interlocutores intencionalmente direcionam a atenção uns dos outros para um aspecto da experiência em detrimento de outros (Tomasello, 2000b: 38), o que possibilita não apenas a aquisição da língua e cultura pela criança mas também o seu desenvolvimento subseqüente em indivíduos adultos.

4.6. O potencial universal das unidades entoacionais

Se, de fato, a prosódia (juntamente com vários outros recursos gestuais) possui uma função básica de realizar enquadramentos atencionais sobre entidades do ambiente

⁹⁶ Essa analogia é sugerida em Tomasello e Farrar (1986: 1462) e meu propósito aqui, ao articular essa proposta com a da análise da conversa, é o de argumentar a favor da similaridade entre objetos lingüísticos e objetos físicos no que concerne aos processos atencionais. O estudo de LSs como a libras oferece evidências interessantes dessa similaridade, como será demonstrado nas análises das seções 7.4.1 e 7.4.2.

imediate e/ou do próprio universo do discurso, o que é crucial para a coordenação das ações sociais, é plausível especular que todas as línguas humanas possuam UEs com função similar. Seria interessante, então, investigar quais das características formais dessas unidades se mostram universais e quais se mostram particulares de uma língua, a fim de dimensionar a possível relevância dessas características para a análise de uma língua tão distinta do inglês – que constituiu a base do estudo de Chafe –, como é o caso da libras.

Não por acaso, desde os trabalhos de Chafe, uma série de estudos sobre UEs tem sido conduzida em várias línguas, com o intuito de aprofundar o conhecimento sobre a natureza e função dessas unidades básicas da fala. Tais estudos, ainda que não muito numerosos, têm tomado como objeto línguas de famílias não-relacionadas tão diversas quanto: inglês e alemão, na Europa; japonês, chinês, coreano e cebuano, na Ásia; hebraico, no Oriente Médio; sêneca, na América do Norte; wardaman, na Oceania; entre outras (Amir et al., 2004; Chafe, 1994; Croft, 1995, 2005; Iwasaki e Tao, 1993; Izre’el, 2005; Matsumoto, 2000, 2003; Park, 2002; Tanangkingsing, 2006; Tao, 1996).

A primeira constatação que podemos apontar nesses estudos – ainda que possa parecer óbvia – é a de que, em todas as línguas, UEs puderam ser identificadas. De maneira geral, o procedimento de segmentação dessas unidades é feito de maneira perceptual e, em alguns casos, recorre-se à análise acústica como complementação para as investigações e para eventuais ajustes nas fronteiras definidas (procedimentos descritos em Du Bois et. al., 1992; ver também Schuetze-Coburn et al., 1991). Para essa identificação perceptual, as pistas prosódicas apontadas por Chafe, além do contorno entoacional, são as pausas, as redefinições de pitch, as mudanças de qualidade da voz e o padrão de aceleração-desaceleração.

Dentre essas pistas, Chafe (1994) aponta o padrão de aceleração-desaceleração como o mais consistente (p. 59), ainda que as pausas sejam os mais salientes (p. 57). Amir et. al (2004), corroborando o achado de Chafe num estudo comparativo sobre os correlatos acústicos e perceptuais das UEs no hebraico, identificam o alongamento final como o principal critério (81% das UEs), seguido da redefinição de pitch (70%), das pausas (55%), e da aceleração inicial (34%). A maioria dos estudos, contudo, têm outros focos de análise e se limita apenas a mencionar os critérios de delimitação adotados.

Além das características formais internas e externas das UEs que permitem a sua delimitação, Chafe enumera ainda outros aspectos importantes dessas unidades que merecem ter o seu potencial universal verificado. Um primeiro aspecto é o tamanho reduzido das UEs, que Chafe analisa por meio do número médio de palavras que essas unidades apresentam. Para as UEs substantivas, Chafe observou uma média de quase cinco palavras, ao passo que para as regulatórias, a média cai para cerca de duas palavras. Dos estudos pesquisados que se ocuparam com essa questão, todos se concentraram na análise das chamadas “UEs substantivas”. Tanangkingsing (2006), no estudo do cebuano, uma língua filipina, identifica a média de quase quatro palavras/UE, sendo que 75,7% das UEs apresentam entre uma e cinco palavras – embora haja UEs de até quinze palavras nessa língua. O autor menciona ainda que a média das UEs no cebuano está muito próxima da do chinês, que oscila entre três e quatro palavras/UE. No hebraico, Amir et al. (2004) identificam uma média de três palavras/UE, a variação absoluta ocorrendo entre uma e seis palavras. Embora seja óbvio que o número de palavras por UE seja um critério bastante relativo, uma vez que línguas tipologicamente distintas do ponto de vista morfológico apresentarão padrões distintos de distribuição de informações em palavras, ainda assim uma média relativamente baixa de palavras/UE parece plausível de ser estipulada como universal, considerando-se a diversidade de línguas estudadas até o momento.

Uma outra observação de Chafe relacionada a esse tamanho reduzido das UEs é a alta correlação entre essas unidades e a forma gramatical da *oração*. As UEs do tipo oracional podem apresentar uma idéia relativa a um estado ou evento, bem como aos referentes que deles participam, embora seja possível encontrar UEs formadas exclusivamente por sintagmas nominais (i.e. restritas apenas a referentes). Nos estudos de Chafe, as UEs oracionais constituíram cerca de 60% do corpus estudado. Essa correlação tem sido um dos principais aspectos investigados por outros lingüistas interessados nas características universais dessas unidades.

A grande maioria dos estudos vem corroborar essa alta correlação apontada por Chafe. O estudo de Croft (1995) sobre o inglês aponta 47,8% das UEs como sendo formadas por orações simples, enquanto seu estudo com o wardaman (Croft, 2005) aponta a mesma correlação em 50,3% dos casos. De maneira geral, a correlação entre UEs e unidades gramaticais revela grande sistematicidade em ambos os seus estudos (cerca de 90%), e vários dos casos que fogem a esse padrão se referem a elementos que,

segundo o autor, são independentes e apresentam as suas próprias funções discursivas (e.g. interjeições, certos advérbios, etc).⁹⁷

Park (2002), no estudo do coreano, alcança uma conclusão semelhante. Ele aponta que a correlação estrita entre oração e UE alcança 48,3% das unidades e que a não-correspondência entre as unidades tende a se concentrar em orações complexas – das quais 74% foram produzidas em duas ou mais UEs, enquanto apenas 22% das orações simples foram quebradas em mais de uma UE. Diversos outros estudos parecem reconhecer essa primazia da oração como unidade básica da fala (para uma proposta que caminha especificamente nesse sentido, ver Thompson e Couper-Kuhlen, 2005). Refletindo sobre essa primazia, Langacker (2001) argumenta que o fato de grande parte das UEs corresponderem a orações simples “representa um mapeamento muito natural, especialmente no caso das orações finitas, que incorporam o ancoramento (*grounding*)”⁹⁸ e, desse modo, indicam como o falante e o receptor enxergam o processo realçado em relação às suas próprias circunstâncias” (p. 154).⁹⁹

No que se refere às manifestações formais da salientação no âmbito de uma UE, Chafe (1994) diz que a proeminência se manifesta principalmente pelo uso de nomes

⁹⁷ Cabe destacar que a opção metodológica pela análise de narrativas (espontâneas e semi-espontâneas) no trabalho de Croft tende a favorecer a correlação entre unidades prosódicas e sintáticas em um grau maior do que podemos esperar da conversa espontânea. Isso ocorre porque as narrativas (especialmente quando produzidas fora de contextos conversacionais, como nas “Histórias da Pêra”, as quais Croft utilizou neste estudo) estão significativamente menos submetidas ao nível de interatividade que se observa na conversação cotidiana.

⁹⁸ O conceito de *ancoramento* de Langacker faz referência aos elementos lingüísticos que estabelecem a relação entre o conteúdo semântico e a situação de fala.

⁹⁹ Alguns poucos estudos têm questionado a validade da oração como unidade gramatical básica da língua espontânea. Iwasaki e Tao (1993), em sua análise comparativa entre inglês, japonês e chinês, categorizam seu corpus em termos de UEs oracionais e UEs não-oracionais – assim como faz Tao (1996) em uma análise do chinês. Eles apontam que, diferentemente do inglês, cujas UEs oracionais prevalecem em seu corpus (53,6%), tanto no chinês quanto no japonês essa primazia não se mantém (60,2% das UEs no chinês são não-oracionais, assim como 54,6% das UEs no japonês). Izre’el (2005), numa análise ainda preliminar do hebraico, divide seu corpus nessas mesmas duas categorias e conclui, assim como os pesquisadores do japonês e do chinês, que a ligeira predominância das UEs não-oracionais (cerca de 52%) em relação às oracionais (cerca de 47%) revelam uma supervalorização da oração nos estudos tradicionais do hebraico.

Esses poucos estudos, porém, parecem fragilizados diante de certas críticas. Croft (2005) identifica problemas metodológicos na conclusão de Tao (1996) que podem ser estendidos aos trabalhos de Iwasaki e Tao (1993) e Izre’el (2005). Croft considera um erro o uso de “UE não-oracional” como categoria analítica, tendo em vista a extrema heterogeneidade de estruturas que ela pode abarcar; em contraposição à sólida coesão da categoria “UE oracional”. Em face dessa discrepância, e considerando que na maioria dos estudos as orações simples constituem algo em torno da metade das UEs, Croft argumenta ser contraproduzitiva a conclusão de Tao, que nega a primazia da oração no chinês falado.

plenos (em contraposição aos pronomes) e de acentos primários (em contraposição a acentos secundários). Em sua comparação com o sêneca, porém, ele já deixa claro que a proeminência é realizada de diferentes maneiras em cada língua, de acordo com as suas características peculiares (ver também, e.g., Lambrecht, 1994). No inglês indiano, por exemplo, o acento e as mudanças de pitch apresentam um padrão marcadamente distinto do inglês ocidental, em especial no que concerne à sua distribuição no enunciado (e.g. na realização de contrastes, o pitch marcado ocorre em orações inteiras ao invés de aparecer localizado em palavras específicas) (Gumperz, 1982). Dentre os recursos de salientação de informação frequentemente discutidos na literatura, então, encontram-se desde sinais prosódicos (e.g. variações de pitch, alongamentos, aumento de volume) até morfossintáticos (e.g. ordem das palavras, construções sintáticas marcadas, morfemas de focalização).

Relacionada a essa discussão sobre a salientação de informações está a tese de Chafe (1980) de que a produção do discurso opere sob restrições de ordem cognitiva, formulada pelo autor a partir da análise dos pequenos “pedaços” de discurso – em geral, orações – que aparecem correlacionados às UEs no discurso. DuBois (1987) estende essa tese buscando especificar, em termos de manifestações lingüísticas observáveis, que tipo de restrições se mostram correlacionadas ao processo de produção do discurso por meio de UEs.

Com base na análise de narrativas numa língua maia, o sacapulteco, DuBois identifica quatro restrições formais relativas à produção de orações que, tomadas em conjunto, constituem o que ele chamou de *estrutura argumental preferida*.¹⁰⁰ No âmbito pragmático, referente ao estatuto informacional de referentes, DuBois propõe que sujeitos transitivos sejam predominantemente dados e que a fala prossiga em termos de não mais do que um argumento novo por UE. No âmbito gramatical, o autor propõe que os falantes evitem sujeitos transitivos lexicalizados e mais do que um único argumento lexicalizado por oração. Tal proposta desencadeou uma série de estudos translingüísticos buscando verificar a sua aplicabilidade universal, o que até agora tem sido demonstrado de maneira consistente (e.g. Ashby e Bentivoglio, 1993, com o francês e o espanhol; Smith, 1996, com o hebraico; Kärkkäinen, 1996, com o inglês americano;

¹⁰⁰ DuBois analisa a estrutura argumental da oração em termos da relação entre o verbo e seus sintagmas nominais nucleares (i.e. seus argumentos), que podem ser do tipo S (sujeito intransitivo), A (sujeito transitivo) e O (objeto direto).

Matsumoto, 2000, com o japonês; Pezatti, 1996, com o português brasileiro, entre outros).

A proposta de que deva haver uma restrição operante no processo de produção da fala – que força as UEs a serem formadas por constituintes relativamente simples (Chafe, 1994), caracterizados por restrições pragmáticas e gramaticais (DuBois, 1987) – é fortalecida nos estudos de Clark e Wasow (1998) e Croft (1995; 2005). Clark e Wasow trazem evidências robustas de que o esforço de expressar uma grande carga informacional a cada novo momento – carga essa avaliada em termos da complexidade sintática de uma fala expressa por uma UE – seja um dos principais fatores preditivos de disfluências na fala.

Numa das demonstrações dessa hipótese, os autores analisam a repetição de artigos sob diferentes contextos sintáticos num corpus de fala espontânea. Tendo em vista que artigos como “the” do inglês aparecem no início de constituintes com diferentes ordens de complexidade, a hipótese é a de que disfluências envolvendo repetições de artigos seriam maiores diante de constituintes complexos do que de constituintes simples. Confirmando essa predição, a pesquisa mostra que os artigos que antecedem NPs complexos (e.g. “the, the time we were there at the warehouse”) são significativamente mais repetidos do que os que antecedem NPs simples (e.g. “the, the diesel”), assim como os artigos que antecedem constituintes maiores (uma oração ou sentença) são mais repetidos do que os que antecedem constituintes menores (objeto do verbo ou da preposição). Análises similares com diferentes tipos de pronomes revelaram essa mesma sistematicidade.

Em seus estudos do inglês e do wardaman, Croft (1995, 2005) adota um olhar inverso ao de Chafe: ao invés de observar quais são as características dos constituintes que aparecem expressos numa UE, ele analisa de que maneira as diferentes construções sintáticas são configuradas em termos de UEs (para análises semelhantes, ver Ono e Thompson, 1995 e Park, 2002). Esse olhar permite ao autor observar, em primeiro lugar, que há uma quebra sistemática de unidades gramaticais complexas em unidades menores, que acabam distribuídas ao longo de dois ou mais agrupamentos prosódicos sucessivos. A partir daí, analisando um número reduzido de unidades complexas que rompem com esse padrão, aparecendo freqüentemente sob uma única UE, o autor nota que tais construções envolvem relações semânticas apresentadas em vários estudos como universalmente propensas à gramaticalização (e.g. “manipulação do sujeito +

[instrumento] + ação”, como em “Ele pegou o martelo e quebrou a janela” que tendem a resultar em construções seriais do tipo “Ele pegou e quebrou a janela”).¹⁰¹

De maneira geral, portanto, os estudos centrados na análise das UEs no discurso espontâneo revelam claras restrições no nível de complexidade das construções que aparecem agrupadas por essas unidades. Permanece carente de maior aprofundamento quais seriam *o(s) fator(es) relevante(s)* que, quando multiplicado(s), parece(m) estourar a nossa capacidade de formulação a cada novo momento, impondo a cisão de uma dada conceitualização ao longo de mais de um quadro atencional. Para os fins deste trabalho, porém, uma resposta definitiva para essa questão não é essencial, uma vez que, no que diz respeito à busca de critérios para a segmentação da conversação espontânea na libras, as características formais das UEs – identificadas por Chafe e expandidas por vários pesquisadores no estudo de outras línguas humanas – oferecem subsídios suficientes para uma investigação inicial sobre a sua aplicabilidade na libras.

4.7. Conclusão

O presente capítulo discutiu um conceito-chave neste trabalho: as UEs. Tais unidades – referidas em alguns trabalhos como “unidades entoacionais”, “enunciados”, entre outras denominações – tiveram suas características investigadas por um grande número de lingüistas de diferentes orientações, embora a sua função, tal como aqui compreendida, tenha sido inicialmente desenvolvida nos trabalhos de orientação funcionalista.

De acordo com os trabalhos pioneiros de Chafe (1994), as UEs refletiriam um foco de ativação de informação que, tendo em vista as nossas limitações cognitivas, viabilizariam a troca de informações entre os interlocutores, salientando um pedaço da experiência humana em detrimento de outros. Langacker (2001) reformula essa função argumentando que, ao realizar esse papel de salientação, as UEs funcionariam como quadros atencionais necessários à coordenação da interação. Dando sustentação a essas propostas, trabalhos desse e de outros autores têm demonstrado que as UEs são apenas

¹⁰¹ Essas análises levam o autor à “hipótese do armazenamento da UE”, segundo a qual as construções gramaticais que aparecem num único agrupamento prosódico tendam a ser processadas como um todo, com estatuto de unidade, ao passo que os constituintes distribuídos em diferentes agrupamentos devam ser computados (Croft, 1995: 872). Nesse sentido, a realização dessas construções complexas em uma única UE seria o primeiro passo formal no processo de gramaticalização dessas unidades, que sofreriam uma espécie de *bootstrapping* sintático por meio da prosódia.

um dos vários níveis de estruturação da língua que manifestam o princípio organizacional de figura-e-fundo.

Além disso, procurei trazer evidências de que esse princípio opera também em níveis sociais mais macro-analíticos, tais como o estabelecimento do “encontro social”, discutido por Goffman (1964); o trabalho ad-hoc de “categorização social”, discutido por Sacks (1972); o funcionamento do um-de-cada-vez na conversação, discutido por SSJ; e a emergência do cisma conversacional, discutido por Egbert (1997). A consistência com que o fenômeno de salientação se manifesta nesses diferentes trabalhos sugere que a operacionalidade da língua e da interação seja fortemente dependente desse que parece ser um princípio básico de estruturação da experiência humana.

A proposta das UEs como unidades básicas do discurso e a sua função na coordenação de ações sociais ganha força também com os estudos de Tomasello e colegas, que mostram que o estabelecimento da atenção conjunta – por meio da coordenação da atenção sobre um único “objeto” salientado – é um pré-requisito fundamental para a aquisição das línguas humanas. Esse processo de gerenciamento da atenção social revela ainda que recursos gestuais diversos são tão importantes quanto a prosódia, questão possivelmente posta de lado nos estudos lingüísticos pelo forte viés em relação aos aspectos vocais da produção.

Diante dessas contribuições para a presente pesquisa, o principal problema que cerca os estudos trazidos nesta seção é o olhar essencialmente mentalista sobre língua e discurso. Nas vertentes lingüísticas aqui estudadas, a gramática continua a ser vista como um sistema que permanece independente do domínio de ação social. Nessa perspectiva, a função atencional das unidades mínimas passa a ser a necessidade de “fazer a ponte” entre os “conteúdos” de “duas mentes individuais” (Chafe, 1994). A AC oferece uma leitura alternativa importante, nesse sentido, na medida em que a função atencional dessas unidades é deslocada das “mentes individuais” para a “ação social”, e da necessidade de “fazer a ponte” para a de “coordenar ações” a fim de atingir determinados propósitos práticos dos interlocutores.

De maneira geral, entendo que a AC ofereça um olhar teórica e empiricamente mais pertinente para abordagem da língua em uso, mas que os estudos lingüísticos – em especial os de orientação funcionalista e cognitivista – especifiquem tecnicamente, de

uma maneira mais rigorosa, alguns conceitos operacionais necessários para o trabalho com os dados propriamente dito.

5. Metodologia: O desenvolvimento do corpus

5.1. Introdução

As metodologias de pesquisa no campo da lingüística estão intimamente associadas aos tipos de dados privilegiados para análise. Chafe (1994) destaca, nesse sentido, duas oposições principais que podem servir de base para o início desta reflexão metodológica: entre dados *públicos* e *privados* (se os dados estão acessíveis a qualquer um que quiser observá-los, ou se estão restritos a um único observador-experienciador) e entre dados *manipulados* e *naturais* (se os dados são estimulados e arranjados pelo observador ou se ocorrem espontaneamente).

Várias das metodologias clássicas de análise lingüística estão associadas às diferentes combinações desses parâmetros, como mostra a tabela 1 abaixo, reproduzida de Chafe (1994: 18):

Dados	Públicos	Privados
Manipulados	Eliciação	Julgamento de gramaticalidade
	Experimentação	Julgamentos semânticos
Naturais	Etnografia da fala	Fluxo do pensamento
	Análise de corpus	Observação introspectiva

Tabela 1. Tipos de dados e principais metodologias de análise lingüística segundo Chafe.

Dados *públicos* e *manipulados*¹⁰² são resultado de experiências e eliciações lingüísticas levadas a cabo por pesquisadores junto a informantes bilíngües a fim de buscar evidências para as questões analíticas específicas que os ocupam. Tal metodologia traz a vantagem de oferecer dados diretamente relevantes à questão do pesquisador, mas traz também como desvantagem o fato de esses dados, afetados por determinadas condições experimentais, muitas vezes distorcerem ou não refletirem a produção lingüística genuína dos falantes da língua.

Dados *públicos* e *naturais*, em contrapartida, envolvem o registro da produção dos falantes em situações de uso espontâneo da língua, por meio de câmeras e

¹⁰² A palavra *manipulado* tem uma conotação pejorativa no português que não deve ser confundida com o uso do termo “dados manipulados” nesta discussão, que pode ser mais bem compreendido em referência a dados *fabricados, direcionados, controlados* experimentalmente.

gravadores previamente dispostos. A vantagem, neste caso, é a de que a língua se apresenta em uma forma mais genuína, tal como os falantes a utilizam, e não como eles próprios ou o pesquisador *acreditam* que ela seja utilizada; a desvantagem, porém, é a de que questões de pesquisa de interesse mais direto do pesquisador deverão aparecer nos dados de maneira acidental, podendo inclusive não aparecer.

Dados *privados e manipulados* dizem respeito aos julgamentos semânticos e gramaticais feitos pelo pesquisador a partir de frases construídas, o que exige a evocação de contextos imaginados que dêem significação a essas frases a fim de que a sua adequação gramatical ou semântica possa ser avaliada. Aqui, novamente, a vantagem para o pesquisador é a de ter uma questão pontual de seu interesse imediato refletida nos dados produzidos, com a desvantagem da possível impertinência desses construtos em relação ao uso espontâneo da língua.

Por fim, dados *privados e naturais* são os mais difíceis de serem obtidos e analisados, uma vez que envolvem aquilo que se passa na mente de forma natural. Embora não haja metodologias específicas voltadas para o uso desse tipo de dado, é evidente que muitas questões de análise emergem inicialmente a partir de introspecções espontâneas, que podem ser, então, posteriormente verificadas em pesquisas com dados empíricos.

Ao romper com a dicotomia entre língua e discurso, ambas as abordagens teóricas que orientam o presente estudo privilegiam os dados públicos e naturais. Mais do que isso, dentre os diferentes tipos de uso espontâneo da língua, os pesquisadores da AC, em particular, têm conferido um estatuto privilegiado à conversação em relação a outras formas de fala-em-interação, tais como debates, palestras, sermões, etc – ver também Chafe (1994: 5) para proposta semelhante.

Como argumenta Schegloff (1996: 54), o cenário primordial da vida social é aquele em que dois seres da mesma espécie interagem face-a-face, isto é, em co-presença física. No caso específico dos humanos, a fala-em-interação parece ser uma das formas mais significativas deste elemento da vida social, e a conversa cotidiana a sua forma básica de organização (p. 54). Assim, apesar de novas tecnologias estarem constantemente gerando cenários sociais inéditos de interação, e conseqüentemente, novos tipos de fala-em-interação que adquirem um grau crescente de importância em sociedades complexas como a nossa, a conversa cotidiana é possivelmente o único

gênero de fala universal a todas as sociedades humanas, além de se constituir no berço para a aquisição da língua pelas crianças (Clark, 2000: 54).

Há razões fortes para crermos, portanto, que as características básicas das línguas naturais foram moldadas de acordo com esse ambiente primordial, como adaptações à, ou como parte da textura que constitui esse gênero de fala específico (Schegloff, 1989: 142-4).¹⁰³ Essa ênfase na conversação, a meu ver, se justifica ainda mais no caso de comunidades como as dos surdos, que, por não possuírem uma escrita para as LSs, estabelecem as suas trocas sociais primordialmente em situações espontâneas próprias da (corp)oralidade (McCleary, 2003).

A presente pesquisa, portanto, privilegia não apenas a análise de dados públicos e naturais, mas, em particular, da conversação espontânea como um gênero de fala primordial. Apesar disso, obstáculos técnicos importantes se colocam para o registro de conversas espontâneas numa língua de natureza gestual-visual, como é o caso da libras. Como discutido no capítulo 2, são amplas as evidências de que as LSs se aproveitam de vários articuladores não-manuais para a realização de funções gramaticais e prosódicas. Isso exige – principalmente no estado incipiente de conhecimento em que ainda nos encontramos – cuidados específicos para garantir que as gravações capturem esses diferentes articuladores com qualidade suficiente para serem descritos.

Por esse motivo, optou-se nesta pesquisa pela utilização de dados públicos que envolvessem tanto uma dimensão *natural* quanto *manipulada*, ainda que a ênfase esteja sobre o primeiro aspecto. Tal combinação, referida como dados de natureza “quase-espontânea” ou “semi-espontânea”, envolve basicamente duas circunstâncias distintas. Uma opção consiste no registro de situações de uso que, a despeito de serem altamente

¹⁰³ Tem-se discutido muito, especialmente em abordagens voltadas para a aquisição de língua e para o discurso, o papel primordial das *narrativas* na constituição dos seres humanos como indivíduos e/ou como povos. A meu ver, contudo, o ato de narrar pressupõe o estabelecimento prévio de uma relação interpessoal em que no mínimo dois participantes se ratificam como mantenedores de um mesmo foco de atenção cognitiva e social – uma conversação, ou nos termos de Goffman (1981), um *encontro*. De fato, como discutido no capítulo 3, as narrativas freqüentemente se manifestam imbricadas nas conversações, sendo licenciadas por meio de estratégias lingüísticas e interacionais específicas por meio das quais os participantes conseguem suspender a relevância dos possíveis pontos para transição entre falantes, a fim de alocar turnos maiores de participação em que a narrativa possa ser levada a cabo.

Vale ainda assinalar que ‘narrativa’, dentro da perspectiva aqui adotada, tem um sentido mais micro-analítico, de uma prática social estruturada para fins específicos numa dada conversação, e não no sentido mais macro-analítico, como um conjunto de textos que refletem, ou oficialmente visam a representar, a ‘cultura’, a ‘ideologia’ ou a ‘identidade’ de um indivíduo ou povo.

direcionadas para uma questão de análise particular, são estrategicamente inseridas em contextos cotidianos a fim de garantir a sua espontaneidade. É o caso, por exemplo, da narrativa elaborada por Moreira (2007) para levantar questões particulares de deixis na libras que, tendo sido previamente trabalhada junto a uma professora surda, foi então inserida numa aula de contação de histórias para crianças de uma escola de surdos (ver também, e.g. Duncan, 2005, para estratégia similar).

Outra opção é o registro de situações de uso da língua que, a despeito de serem eliciadas pelo pesquisador, são pouco ou nada direcionadas por questões particulares de pesquisa. É o caso da eliciação de narrativas e conversas entre surdos em estúdios de gravação, sem tópicos ou temas previamente definidos, que serviram de base para algumas análises em Liddell (2003a).¹⁰⁴ No caso da presente pesquisa, então, optou-se por essa estratégia menos diretiva, isto é, pela gravação de conversas entre amigos surdos num estúdio previamente arranjado.

Privilegiar dados mais naturais (corpus espontâneo ou quase-espontâneo) em detrimento de dados mais manipulados (sentenças eliciadas ou construídas com fins de julgamentos de gramaticalidade) tem sido uma estratégia crescente entre vários lingüistas, em especial aqueles com orientação teórica voltada ao uso. Pesquisadores como Hopper (2001) chegam a especular que a revolução na natureza dos dados trazidos por tecnologias emergentes de registro e manipulação de dados espontâneos, principalmente a partir da década de 90, irá se constituir como o grande divisor de águas nos estudos lingüísticos. O que está em debate, mais do que uma questão de “preferência”, é o modo como o tipo de dado que serve de base para as pesquisas pode conduzir os lingüistas a análises bastante distintas sobre um mesmo fenômeno.

Em sua análise sobre as sentenças pseudo-clivadas a partir de um corpus de discurso espontâneo, por exemplo, Hopper (2001) questiona, com base em seus achados, o argumento de que os dados ao qual as crianças em fase de aquisição estariam expostas seriam uma “evidência degenerada” da língua:

Corpus studies suggest ... that these “degenerate” data are the true substance of natural spoken language, and that what our descriptive and prescriptive grammars give us are normativized assemblies of these fragments that tend to impress themselves on us as mental prototypes because of their greater social prestige – their associations with

¹⁰⁴ Apesar disso, Liddell (2003a) não explicita a sua metodologia de trabalho, e, em algumas partes do livro, a sua análise parece também ser baseada em sentenças eliciadas.

schooling, with literacy, and with complex discourse characterized by long periods and uninterrupted turns. This observation has significant consequences for both linguistic theory and applied linguistics.

Se esse alto grau normativo das construções gramaticais canônicas é um ponto crítico na análise de LOs, ainda mais crítica é sua utilização como ponto de partida para a análise de línguas sem escrita, como é o caso das LSs. Nesse contexto, o apoio sobre a “gramática” das línguas ocidentais – com todo o seu histórico de hegemonia da escrita – põe em risco ainda mais a possibilidade de identificação da “real substância” da língua. Como alerta Hopper (2001):

By the time a grammatical construction comes to the attention of linguists it is already highly normativized. Worse still, Westernized normative standards may be smuggled into descriptions of unwritten languages when linguists base their elicitation on English equivalents. To view the canonical constructions as prototypes and as the source of ‘deviant’ fragmentary instantiations in discourse is to put the cart before the horse (p. 126).

Tais considerações não implicam que estratégias de produção de dados manipulados não tenham valor nos estudos lingüísticos.¹⁰⁵ O que se sugere é a necessidade de inverter a ordem de prioridades, partindo-se sempre da observação de dados naturais como principal fonte de levantamento e verificação de hipóteses, e utilizando-se de dados manipulados (i.e. experimentações, eliciações e julgamentos de gramaticalidade) como estratégias complementares para o aprofundamento de questões levantadas. Diante do estágio ainda incipiente do trabalho de descrição gramatical da línguas, então, a formação do corpus e a análise desta pesquisa se restringiram aos dados naturais, reservando-se para um momento futuro a eventual necessidade/possibilidade de combinação com dados manipulados.

Nas seções a seguir, então, serão discutidos em maiores detalhes alguns aspectos metodológicos que cercaram a pesquisa como um todo, desde o processo de gravação das conversas, passando pela necessidade e/ou utilidade de manipulação dos vídeos, e chegando até os recursos e convenções adotados na transcrição dos dados e na produção de relatórios sobre as análises. Nessa reflexão, serão discutidas tanto as dificuldades quanto as soluções que foram encontradas ao longo da pesquisa. O caráter inicial deste

¹⁰⁵ Nesse sentido, Labov (1996) oferece uma rica discussão sobre como, no desenvolvimento de estratégias de manipulação de dados, o pesquisador pode evitar determinadas circunstâncias que conduzam a erros de intuição dos falantes e/ou dele próprio.

tipo de estudo no Brasil revela, porém, que soluções metodológicas mais robustas para essas questões somente serão alcançadas com o avanço dos trabalhos lingüísticos com base em corpus informatizado da libras.

5.2. Gravação das conversas

Para a etapa de gravação, três surdos adultos, fluentes em libras, foram convidados para participar voluntariamente do estudo.¹⁰⁶ Os participantes se revezaram em três duplas, que foram deixadas a sós por 20 minutos no estúdio de gravação do Centro de Comunicação Eletrônica (CCE) da USP para conversar sem nenhum tópico previamente definido. Dentre elas, duas duplas eram compostas por amigos que já se conheciam há anos, ao passo que uma era formada por conhecidos com pouca intimidade. Aos participantes foi informado apenas que o pesquisador estava interessado em analisar o uso espontâneo da libras entre surdos.

No estúdio, cinco câmeras haviam sido previamente preparadas de modo a captar diferentes tomadas dos participantes: duas tomadas para captação do espaço de sinalização¹⁰⁷ referente a cada participante; duas tomadas para captação do rosto de cada participante; e uma tomada para captação da ação conjunta dos dois participantes em perfil. Com a colaboração da equipe do estúdio multimeios do CCE, esse aparato pôde ser implementado e a gravação realizada, tal como mostra a figura 11 abaixo.

A exigência de todo esse equipamento, ao mesmo tempo em que viabilizou uma análise mais fina da libras, também afastou as condições de gravação do ideal almejado no campo da AC (Schegloff et al., 1996: 16-19). Nesse sentido, três tipos de observação são pertinentes. Em primeiro lugar, o fato de duas das três duplas serem formadas por antigos amigos, a meu ver, se refletiu numa maior espontaneidade de suas conversas, que, diferentemente da dupla cujos membros eram menos familiares entre si, se mostraram altamente relaxadas e descontraídas no desenrolar da conversa. Essa diferença levou-me a optar por utilizar a conversa de uma dessas duplas de amigos como base para a transcrição e análise nesta pesquisa.

¹⁰⁶ Os participantes assinaram em cartório um termo de compromisso autorizando o uso dos dados em pesquisas, e, referência a eles no capítulo de análise será feita pelo uso da inicial de seus nomes: (R)egiane, (W)ilson e (S)andro. Outros nomes particulares que aparecem nas conversas gravadas foram alterados e ocultados nos vídeos a fim de se preservar a identidade de seus referentes.

¹⁰⁷ O espaço de sinalização é, grosso modo, o espaço a frente do sinalizador, incluindo o seu próprio corpo, onde a sinalização tende a ocorrer.

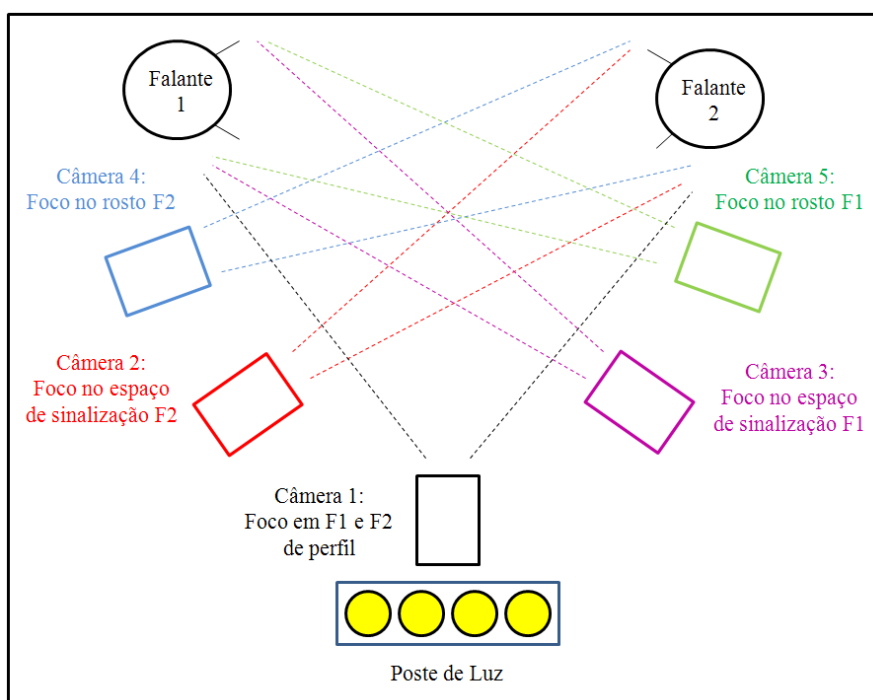


Figura 11. Disposição das câmeras e da iluminação na gravação.

Em segundo lugar, por mais relevante que sejam os esforços de neutralizar os constrangimentos de gravação para os participantes numa pesquisa, mais fundamental à coerência metodológica do estudo é o registro e a explicitação das circunstâncias da gravação, a fim de que a análise possa ser apropriadamente avaliada e dimensionada em face do contexto de emergência dos dados. Cada situação de fala particular, seja menos ou mais espontânea, deve sempre ser compreendida dentro das circunstâncias que a suscitam e que a restringem.

Por fim, como argumentam Heritage e Greatbatch (1991), não é o ambiente da interação em si que irá revelar ao analista uma organização da fala distinta, mas sim a orientação que os próprios participantes demonstram uns em relação aos outros no curso da conversa. Não existe, na perspectiva etnometodológica da AC, uma caracterização analítica da fala (do tipo natural, institucional, ou qualquer outra) que possa ser definida *a priori*, sem uma consideração dos dados empíricos sob investigação (Schegloff, 1997).

No que diz respeito ao aspecto mais técnico da gravação, o uso posterior dos vídeos revelou que a disposição das câmeras e da iluminação não foi a mais adequada. As tomadas referentes ao espaço de sinalização (câmeras 2 e 3), por estarem muito diagonais em relação aos falantes, acabaram dificultando um pouco a visualização de

algumas sinalizações, sempre que o falante se voltava para o lado oposto ao das câmeras. Com relação à iluminação, como mostrado na figura 11, o poste de luz foi originalmente colocado atrás da câmera 1, com foco nos dois participantes de perfil. Como resultado, as demais tomadas, principalmente as com foco nos rostos dos participantes, ficaram muito perpendiculares ao feixe de luz, o que obscureceu de forma significativa a imagem dos participantes. A incerteza sobre movimentos faciais sutis (e.g. piscadas de olhos) acarretada por esse problema de gravação tornou a anotação dos dados mais trabalhosa.

Um outro aspecto que prejudicou a visualização da sinalização na etapa posterior de transcrição dos dados foi a vestimenta. O uso de óculos por um dos participantes em uma das conversas dificultou a observação de movimentos faciais na região dos olhos (sobrancelha, direção do olhar e pálpebras). Tendo em vista que a conversa na qual esse problema foi identificado era também a conversa na qual os participantes possuíam menos intimidade entre si, optei por não utilizá-la como base de análise central neste estudo.

5.3. Manipulação dos vídeos

No processo de transformação da gravação em arquivos que possam ser trabalhados no computador, alguns problemas também tiveram que ser solucionados. O ELAN, programa utilizado para as transcrições nesta tese, suporta especialmente vídeos no formato .mpeg e .mov, mas até mesmo esses formatos podem trazer problemas operacionais caso os vídeos não sejam gravados utilizando as opções mais apropriadas. Foi o que aconteceu numa fase inicial do trabalho, gerando problemas de utilização dos arquivos de dois tipos: certas operacionalidades importantes do ELAN deixaram de funcionar (e.g. a opção de alteração da velocidade do vídeo, que permite fazer a anotação vendo a sinalização em câmera lenta); e outras funcionavam de maneira inadequada, o que atrapalhava ou mesmo impedia o trabalho de transcrição (e.g. imprecisão da sincronização entre o vídeo e as anotações; não abertura da tela de anotação, que permanecia travada já na inicialização do novo arquivo).

Bickford (2005) aponta que os resultados mais satisfatórios em seu trabalho com o ELAN foram conseguidos com a criação de arquivos no formato .mpeg, desde que acompanhados de algumas especificações. De fato, desde que esse artigo foi descoberto

e as sugestões do autor foram adotadas, os problemas de operacionalidade dos arquivos utilizados para transcrição foram resolvidos.

Com os arquivos prontos para a transcrição, então, outras questões pertinentes à manipulação dos vídeos surgiram ao longo da pesquisa. Como discutido na seção anterior, problemas de iluminação causaram o escurecimento de algumas tomadas, principalmente aquelas – com foco no rosto dos participantes – que estavam por demais perpendiculares ao poste de luz. Esse problema foi parcialmente remediado com o tratamento posterior dos vídeos por meio do adição de brilho e contraste, como mostra a figura 12 abaixo.

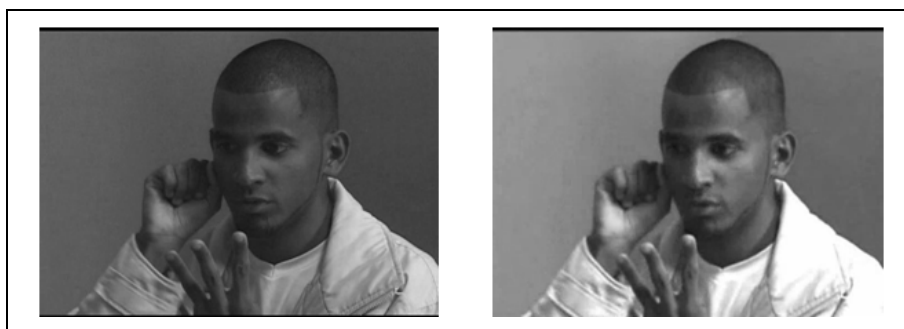


Figura 12. Imagem original gravada e a mesma imagem tratada com auxílio de programa de vídeo.

Na verdade, vários programas permitem esse tipo de manipulação de arquivos de vídeo e o conhecimento desses recursos pode ser útil não apenas para corrigir problemas de gravação, como nesse caso, mas para melhorar ainda mais as condições da etapa de transcrição, no que diz respeito a certas dificuldades inerentes ao processo. É o caso, por exemplo, da anotação e análise que envolvem o deslocamento de mãos, cabeça e tronco do sinalizador no espaço. Em vários momentos da transcrição, tive dificuldades para decidir se o sinalizador de fato havia realizado um dado movimento, ou se uma diferença percebida no deslocamento das mãos no espaço era ou não significativa. Num dado momento, então, descobri que havia a possibilidade de criar um arquivo de imagem contendo uma grade com fundo transparente para ser sobreposta aos arquivos com as gravações das conversas. A figura 13 abaixo ilustra esse recurso, que, na passagem em questão, facilita a apreensão de um movimento sutil da cabeça para trás entre dois sinais consecutivos.¹⁰⁸

¹⁰⁸ Essa estratégia elimina a necessidade de se colocar um fundo contendo uma grade com cores contrastivas atrás dos participantes na etapa de gravação (Scott Liddell, comunicação pessoal).

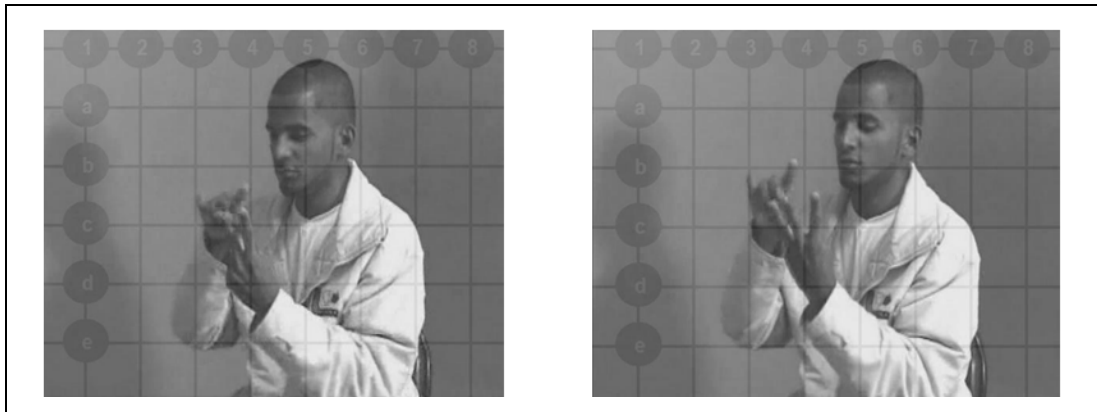


Figura 13. Sobreposição de grade ao vídeo para facilitar anotações de deslocamento do corpo no espaço.

Desse modo, pude produzir dois arquivos de transcrição para cada tomada de gravação relevante à análise dos movimentos: um primeiro vídeo, sem alteração, para transcrição de elementos mais facilmente visualizáveis; e um segundo vídeo, com a grade sobreposta, para a transcrição específica de elementos que precisam ter a sua dinâmica de deslocamento no espaço descrita, como a dinâmica dos sinais manuais e dos movimentos da cabeça e do tronco. Esse recurso acabou se mostrando relevante também para a demonstração de algumas modulações de sinais na discussão da análise, como a seção 6.4.3 irá ilustrar.

5.4. Transcrição dos dados

O ELAN é um programa desenvolvido pelo Max Planck Institute for Psycholinguistics, da Holanda. Entre os vários motivos para adotá-lo em minha pesquisa, cabe destacar: i) a compatibilidade com PCs; ii) a sua distribuição gratuita na internet; iii) a sua crescente utilização em pesquisas com diversas línguas no mundo; iv) o fato de ter sido projetado para viabilizar uma transcrição mais eficiente das LSs; v) as atualizações contínuas e a abertura dos desenvolvedores do programa a sugestões e dúvidas dos usuários; e vi) funcionalidades específicas tal como a sincronização do vídeo com as transcrições, um complexo sistema de buscas, e a capacidade de operar com até quatro câmeras simultaneamente. Todas essas qualidades tornaram o ELAN um programa imprescindível para a viabilização deste estudo.

A estrutura do arquivo de transcrição do ELAN foi basicamente a mesma que foi elaborada pelo grupo “Estudos da Comunidade Surda” (ECS), a partir de uma experiência-piloto de transcrição de uma narrativa dentro do projeto de formação de

corpus da libras (McCleary e Viotti, 2007). Nessa experiência, contudo, a anotação foi feita manualmente, com editor de texto comum, de modo que o meu trabalho foi o de adaptar esse sistema, sob a orientação dos professores do grupo ECS, ao formato do ELAN. Esse programa permite que se criem tantas trilhas (*tiers*) quantas forem necessárias para a anotação dos diferentes articuladores corporais (ver parte de baixo da figura 14) e o anotador tem a opção de trabalhar apenas com aquelas que sejam de seu interesse imediato a cada momento.

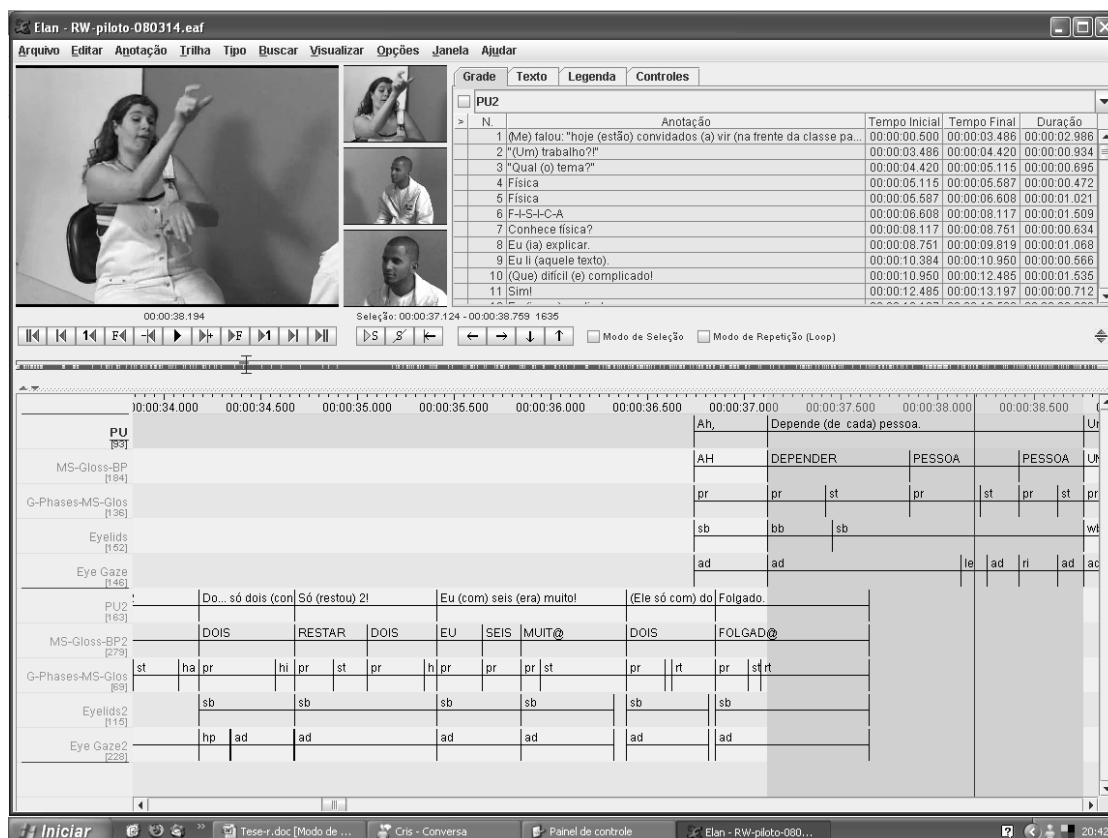


Figura 14. A tela de transcrição do ELAN, com as trilhas anotadas aparecendo na parte de baixo da figura.

O arquivo-modelo completo do ELAN para os trabalhos de transcrição do grupo ECS, que utilizei também em meu trabalho, inclui as trilhas relacionadas na tabela 2 abaixo. As abreviações nos títulos de algumas trilhas são relativas às iniciais do inglês, pois um dos objetivos desse trabalho de formação de corpus é o de torná-lo acessível a pesquisadores de outros países. As trilhas apresentadas estão duplicadas em meu arquivo pela necessidade de transcrever a atividade de ambos os participantes das conversas gravadas. A coluna à direita indica se a trilha em questão foi utilizada nesta pesquisa em vista de meus propósitos atuais (+) ou se não foi considerada (-).

Título da trilha	Descrição da trilha	Utilização
PU	Segmentação do discurso da libras em unidades entoacionais	+
MS-Gloss-BP	Registro de glosas, em português brasileiro, referentes aos sinais manuais	+
MS-Gloss-E	Versão das glosas de MS-Gloss-BP para o inglês	-
NMS-Gloss-BP	Registro de glosas, em português brasileiro, referentes aos sinais <i>não</i> -manuais	+
NMS-Gloss-E	Versão das glosas de NMS-Gloss-BP para o inglês	-
SMS-Gloss-BP	Registro de glosas, em português brasileiro, referentes aos sinais manuais realizados simultaneamente ao sinal glosado em MS-Gloss-BP	+
SMS-Gloss-E	Versão das glosas de SMS-Gloss-BP para o inglês	+
G-phases	Registro das fases dos gestos manuais durante a sinalização	+
Head	Registro dos movimentos de cabeça	+
Eyebrow	Registro dos movimentos de sobrancelha	+
Eyelids	Registro dos movimentos de pálpebra	+
Eyegaze	Registro do direcionamento do olhar	+
Mouth Pictures	Registro das imagens bucais relacionadas aos fonemas do português brasileiro (visemas)	+

Mouth Gestures	Registro dos gestos bucais que não têm correlação com a língua portuguesa	+
Body	Registro de movimentos do tronco	+
Shoulders	Registro de movimentos dos ombros	+
Hands	Registro de qual mão realiza o sinal	+
H-Location	Registro da localização da mão no espaço de sinalização durante a realização do sinal	+
Repetition	Registro do número de vezes que um movimento é repetido no sinal	+
Body Gesture	Registro de gestos não-lingüísticos	-
Comments	Registro de comentários que surgem ao longo da transcrição	+

Tabela 2. Descrição das trilhas que compõem o arquivo de transcrição do ELAN.

A maior parte dessas trilhas é constituída por um repertório fechado e/ou restrito de possibilidades de anotação – exceto quando a trilha apresenta uma gama muito grande de possibilidades, como é o caso das glosas. Esses repertórios de entrada, denominados “vocabulário controlado” no ELAN, facilitam o processo de transcrição e, por esse motivo, foram emprestados do projeto VisiCast de Hamburgo, que já possuía um levantamento de sinais não-manuais registrado a partir de análises de corpus da língua de sinais alemã (Hanke, 2000).¹⁰⁹

Com relação à trilha das imagens bucais (*mouth pictures*), as anotações têm sido feitas de acordo com um repertório de visemas, isto é, de “padrões visualmente contrastáveis de movimentação articulatória dos lábios e da língua produzidos em consequência da fala” (McCleary e Viotti, 2007: 21). No caso da libras, cujos falantes estão numa situação de contato lingüístico com falantes do português, os visemas que

¹⁰⁹ Com o desenvolvimento do corpus na libras em trabalhos como este, a identificação de diferenças entre essas duas línguas deverá permitir que alguns desses valores, ausentes na libras, sejam eliminados de nossos arquivos de anotação, da mesma maneira que outros valores não presentes no repertório, específicos da libras, deverão ser incluídos.

nos interessam são os visemas do português brasileiro. Para isso, foi utilizado o repertório de visemas dependentes do contexto fonético levantado no trabalho de De Martino (2005).

A inclusão da trilha sobre as fases do gesto foi motivada por um desejo de evitar assumir, neste momento inicial, qualquer proposta fonológica acerca da estrutura interna dos sinais (e.g. Brentari, 1998; Liddell e Johnson, 1989; Sandler, 1989). Num esforço de segmentar a cadeia gestual da maneira mais objetiva possível, optei por utilizar o conceito de *fases do gesto*, elaborado no trabalho de Kita et al. (1997) a partir da literatura clássica sobre gestualidade (Kendon, 1972, 1980; McNeill, 1992).

Kendon e McNeill identificaram nos gestos que co-ocorrem com a fala – tipicamente aqueles realizados pelas mãos e braços – fases qualitativamente distintas. A partir daí, descreveram uma hierarquia de unidades que constitui a organização interna dos gestos e que foi estendida por Kita et al. para abranger não apenas os gestos que co-ocorrem com a fala, mas também a sinalização que ocorre nas LSs.

A organização interna dos gestos para um articulador (manual) pode ser esquematizada tal como mostra a figura 15, adaptada de Kita et al. (1997: 27):¹¹⁰

¹¹⁰ Embora essa apresentação siga a proposta de Kita et al. no que diz respeito à definição de conceitos e organização da hierarquia, optei por modificar parte da terminologia dos autores, que se referem a “*movement units*, *movement phrases* e *movement phases*”, em favor de uma terminologia mais próxima de McNeill, que privilegia o qualificativo *gesture*. A meu ver, a idéia de *gesture* é mais abrangente do que a idéia de *movement*, uma vez que o movimento é apenas parte de um todo gestual que envolve, entre outras coisas, pontos de articulação e configurações das mãos. Tal consideração me motivou a adotar as expressões “gestual” ou “do gesto” para qualificar os elementos em diferentes níveis da hierarquia: unidade (*unity*), seqüência (*phrase*) e fase (*phase*).

Vale a pena destacar, mais uma vez, o estatuto essencialmente descritivo desses conceitos. Em particular, as referências a *movimentos* e *suspensões* não devem ser confundidas com as noções de movimentos e suspensões de estatuto fonológico, trabalhadas em teorias sobre a fonologia das LSs tais como a de Liddell (1984), citada na seção 2.2. Em última instância, é possível que haja correlações entre o que aqui se descreve com enfoque empírico e o que se tem proposto sobre a fonologia das LSs. Essa é, contudo, uma questão a ser investigada futuramente e que não entra no escopo deste trabalho.

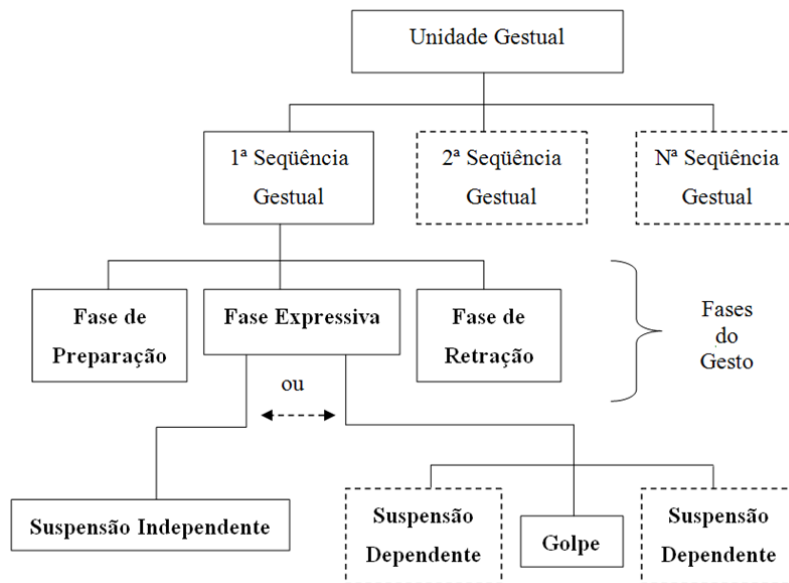


Figura 15. Hierarquia dos elementos que compõem a estrutura interna de gestos e sinais.

A *unidade gestual* envolve os períodos de atividade das mãos, desde o momento em que elas saem da posição de repouso, até o momento em que elas retornam. Durante esse período, as mãos podem realizar, no mínimo, uma *seqüência gestual* – os quadros pontilhados indicam a não-obrigatoriedade das demais seqüências –, que é formada por três fases: i) na fase de *preparação*, os braços e mãos se preparam para a realização da fase expressiva; ii) na *fase expressiva*, que carrega o significado do gesto, as mãos e braços podem ficar suspensos no ar por alguns instantes (a chamada *suspensão independente*) ou as mãos e braços podem realizar um pico de esforço gestual (o chamado *golpe*), opcionalmente precedido e sucedido de suspensões das mãos e braços no ar; e, por fim, iii) na fase de *retração*, as mãos retornam em direção ao repouso, parcial ou totalmente.

Nas LSs, cada sinal corresponde a uma *seqüência gestual* que, abstraída de contexto, sempre apresenta as três fases de movimento acima mencionadas. Na figura 16 abaixo, apresento uma ilustração da estrutura interna do sinal EXPLICAR da libras.

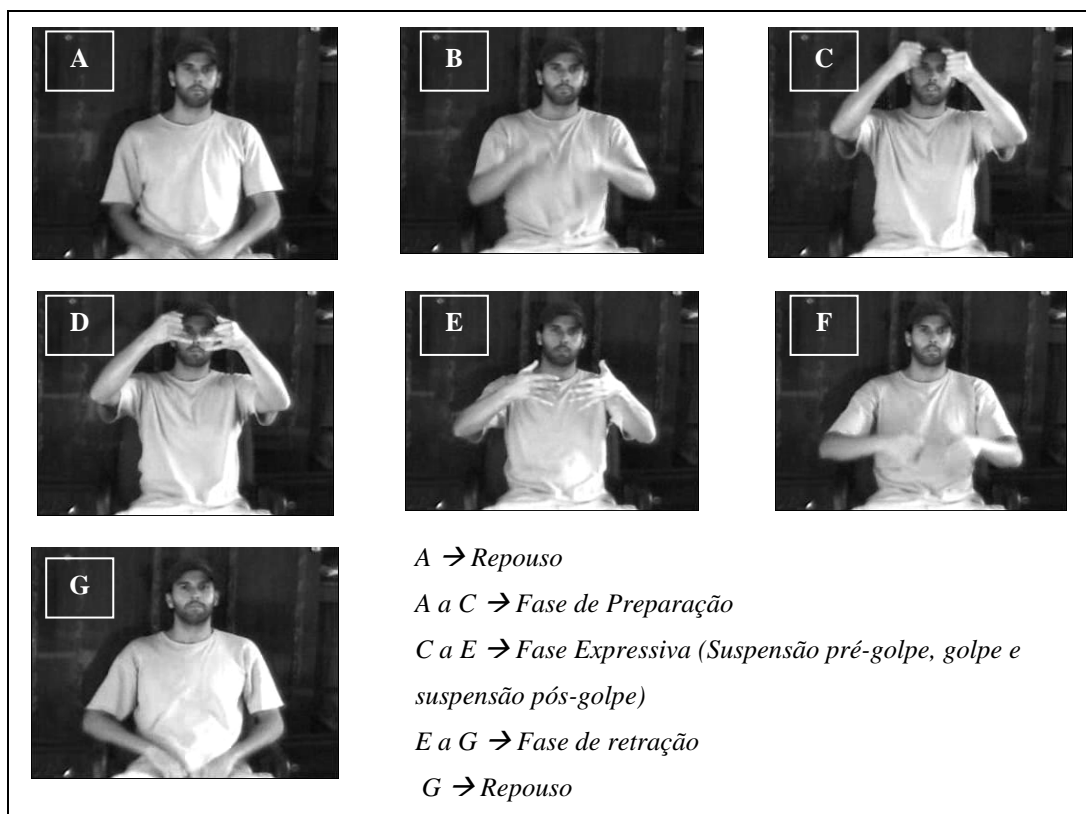


Figura 16. Fases do gesto para o sinal EXPLICAR, cuja fase expressiva é formada por golpe.

A outra possibilidade de realização de uma seqüência gestual (fora de contexto) ocorre quando a fase expressiva é formada *não* por um golpe, mas sim por uma única suspensão independente. Na figura 17 abaixo, apresento uma ilustração da estrutura interna do sinal CINCO da libras, para mostrar essa segunda possibilidade de manifestação da fase expressiva.

Uma diferença dessas descrições em relação às seqüências gestuais em contexto, isto é, dentro da cadeia de fala na libras, é a de que as únicas fases que parecem ser obrigatórias na fala espontânea – pelas observações feitas até o momento – são a fase de preparação e a fase expressiva. A fase de retração, diferentemente, pode ser elidida em alguns contextos ou realizada apenas parcialmente (quando as mãos e braços são mantidas suspensas no ar com um nítido relaxamento da localização e da configuração de mão do sinal, sem retornar de fato a uma posição de repouso). Outra diferença é a de que a seqüência gestual fora de contexto coincide com a unidade gestual (que tem o repouso tanto como ponto de partida quanto como ponto de chegada), ao passo que, na fala, essa coincidência é rara, já que dificilmente o sinalizador produz turnos compostos por um único sinal.

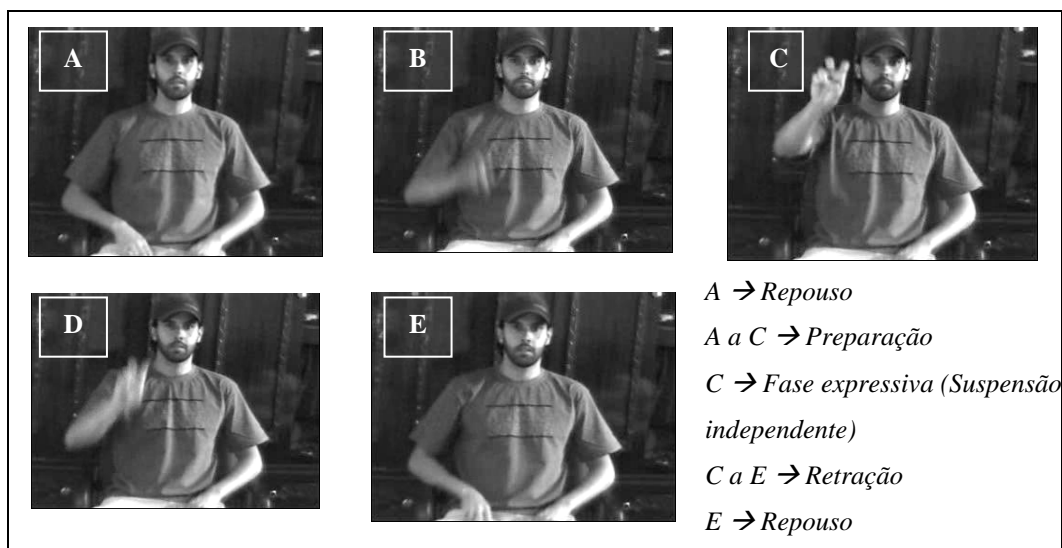


Figura 17. Fases do gesto para o sinal CINCO, cuja fase expressiva é formada por suspensão independente.

Nesta pesquisa, em vista de meus propósitos imediatos, a transcrição não envolveu toda a sofisticação do modelo de Kita et al. Apenas incluí notações no ELAN para as fases que compõem a “seqüência gestual”: i) a preparação; ii) o golpe, a suspensão independente e a suspensão dependente dos golpes, referentes à fase expressiva; e iii) a retração. Além disso, busquei seguir o critério principal de segmentação em fases proposto no estudo de Kita et al., segundo o qual um “movimento é dividido em duas fases se há uma mudança abrupta de direção do movimento da mão *e também* se há descontinuidade no perfil de velocidade do movimento da mão, antes e após a mudança abrupta de direção” (p. 29). Também de acordo com esse estudo, ambos os critérios foram identificados de maneira impressionística, sem um aparato de análise física do deslocamento das mãos no ar.¹¹¹ A relevância desse modelo será ilustrada nas análises sobre a modulação dos sinais, nos capítulos 6 e 7.

Como pode ser observado pelas características das trilhas adotadas, a proposta da transcrição nesta fase da investigação da libras foi a de fazer uma descrição que

¹¹¹ A mudança de direção é bastante fácil de ser apreendida perceptualmente, diferente do que ocorre com a mudança de velocidade. Apesar disso, a segmentação natural dos vídeos digitais em *frames* facilita esse trabalho de identificação de mudanças na velocidade do gesto. Em geral, quando o gesto é mais rápido, o seu congelamento num frame resulta numa espécie de borrão que nos mostra apenas de maneira aproximada a localização exata dos braços ou das mãos no espaço. Diferentemente, quando um gesto é realizado lentamente, o seu congelamento num frame mostra braços e mãos com contornos mais bem definidos no espaço. Esse método impressionístico de análise tem sido utilizado por pesquisadores com longa experiência em estudos da gestualidade (e.g. Duncan, 2005).

evitasse a imposição prematura de categorias lingüísticas (McCleary e Viotti, 2007). Embora seja um fato amplamente constatado que toda transcrição implica, inevitavelmente, uma análise, é igualmente verdadeiro que uma transcrição pode estar imbuída de diferentes graus de motivação teórica, de modo que uma cautela metodológica neste atual estágio de investigação da libras parece prudente.

Ao final, apenas três dos sessenta minutos de conversação gravados foram transcritos, ainda que outras passagens das gravações tenham sido também utilizadas na busca por uma maior generalização de certos fenômenos analisados. A tarefa de transcrição, de fato, é possivelmente a mais difícil de todo o processo, pelo tempo de trabalho que exige do pesquisador. A recompensa, porém, vem não apenas nas diversas hipóteses de análise que vão surgindo a partir das incontáveis repetições no processo de observação dos dados, mas também nos progressos alcançados no desenvolvimento do sistema de notação informatizado como um todo. Trata-se, afinal de contas, de uma primeira etapa necessária que, uma vez superada, viabilizará a formação de um corpus da libras e, conseqüentemente, uma base sólida de dados que poderá servir de subsídio para futuras pesquisas na área.

5.5. Produção de relatórios

Uma das formas utilizadas para apresentação dos dados de análise nesta pesquisa é a proposta elaborada na experiência-piloto do projeto de formação de corpus (McCleary e Viotti, 2007). Nesse trabalho, foi proposto um sistema de notação, que aparece ilustrado na figura 18 abaixo.

/\.....				
V .<.....				
19.	AGORA ₂	<i>FALTAR</i> ₁	OBJETO-REDONDO ₁	RESTAR VAZI@
	2m(1x)	2m(2x)	<2m(1x)<	md(1x) <md(1x)<
				<me BASE-OBJETO.....<
	Agora, falta [o] resto [das pêras] [na] cesta vazia à esquerda			

Figura 18. Sistema de notação proposta em McCleary e Viotti (2007).

A pauta como um todo representa uma (possível) UE na libras, na época segmentada de maneira mais intuitiva, considerando-se basicamente a forma da oração simples. No centro da pauta, em letras maiúsculas, aparecem as glosas referentes aos sinais manuais, que, quando marcadas em itálico, indicam que o sinal se encontra dicionarizado (Capovilla e Raphael, 2001). Marcações morfológicas nominais no português são substituídas pelo sinal de '@' – desde que não comprometam a

inteligibilidade da transcrição – e as marcações verbais são substituídas pela desinência de infinitivo. O hífen que aparece em OBJETO-REDONDO indica que ambas as palavras se referem a uma única glosa, isto é, a um único sinal da libras. Os números subscritos junto às glosas servem para diferenciar sinais que recebem o mesmo “nome” (i.e. a mesma glosa), mas que na verdade apresentam formas diferentes.¹¹²

A organização da pauta é relativamente icônica, tendo em vista que os articuladores do rosto aparecem nas trilhas acima da glosa em ordem similar à que existe no corpo: cabeça, sobrancelha, olhos, boca, ombro, tronco. Na figura 18 acima, apenas as trilhas da sobrancelha e olhos estão anotadas.¹¹³ A primeira trilha refere-se à sobrancelha, que apresenta convenções icônicas tal como o sinal ‘^’, indicador de sobrancelha levantada. A trilha imediatamente abaixo refere-se ao olhar e também envolve sinais que buscam representar de maneira icônica o seu direcionamento, tal como ‘<’, indicador de olhar para a esquerda e ‘|’, indicador de olhar direcionado ao interlocutor.

A trilha abaixo da glosa refere-se às características do sinal manual: se ele é feito com a mão direita (md), esquerda (me) ou as duas mãos (2m); quantas vezes o seu movimento é repetido (1x, 2x, etc); e em que lugar do espaço ele foi realizado (e.g. ‘<’ indica que foi realizado à esquerda). Na trilha logo abaixo, são anotados sinais manuais que co-ocorrem com outros sinais, como no caso em que o narrador sinalizou simultaneamente RESTAR e VAZI@ com a mão direita, e BASE-OBJETO com a mão esquerda. A marcação em cinza, assim como o pontilhado nas trilhas da sobrancelha e do olhar, indica a permanência aproximada do sinal manual ou não-manual em relação à sucessão das glosas. Por fim, a última trilha abaixo da glosa traz a tradução da passagem para o português, destacando entre colchetes elementos da língua portuguesa cuja fonte não provém dos sinais manuais, mas de outros aspectos significativos da libras ou da situação de fala (e.g. o uso icônico do espaço, o contexto da narração).

¹¹² Os nomes dos sinais são mantidos fixos, independente de o contexto onde emergem sugerir um significado incompatível com o nome convencionalmente adotado. Embora os nomes dos sinais sejam dados buscando o máximo de generalidade possível a fim de evitar precisamente tais discrepâncias, a possibilidade desse tipo de incongruência resulta da falta de correlação unívoca entre o léxico da libras e o léxico do português, que serve de recurso para a glosagem.

¹¹³ Por dificuldades técnicas de manipulação do vídeo-cassete, utilizado na época, esse sistema não envolveu a transcrição da cabeça, boca e bochechas, ombro e tronco. Essa segunda etapa está sendo agora desenvolvida com a exploração do ELAN.

Entre os pontos fortes desse sistema de notação está a representação relativamente icônica de vários dos aspectos que compõem a sinalização fluente, o que facilita a sua leitura uma vez que as convenções sejam explicitadas ao leitor; a sua principal fraqueza, porém, está na impossibilidade de uma representação precisa da temporalidade, captada de maneira vaga pelos pontilhados e o espaçamento. Nesse aspecto, a opção de produção de relatórios diretamente do ELAN também não contribui muito, tendo em vista que tais relatórios não preservam a temporalidade das transcrições e as anotações acabam todas justapostas umas às outras em série.

Por esses motivos, nos casos em que a temporalidade for crucial para as análises – como, por exemplo, nas análises da seção 6.4 desta tese – a alternativa adotada foi a de reproduzir diretamente a transcrição do ELAN utilizando o recurso “PrintScreen” do teclado e, em seguida, recortando a parte relevante para análise, como mostra a figura 19 abaixo:

	00:00:47.500	00:00:48.000	00:00:48.500	00:00:49.000	
Eyebrows	rb				
Eye Gaze	dn	ad	be		
MS-Gloss-BP	AGORA2	FALTAR1	OBJETO-REDON	RESTAR	VAZI@
SMS-Gloss-BP					BASE-OBJETO-L
H-Location	L5	L5	L6	L6	L6
Hands	2h	rl	2h	r	r
Repetition	1x	2x	1x	1x	1x
IU	Agora, falta [o] resto [das péras] [na] cesta vazia à esquerda				

Figura 19. Foto tirada da tela do ELAN com a passagem relevante para discussão recortada.

Essa outra forma de apresentação dos dados, porém, apresenta as suas próprias fraquezas. Diferentemente do sistema adotado na pauta acima, as convenções emprestadas do projeto VisiCast que aparecem na transcrição no ELAN não são icônicas, constituindo-se, em geral, de iniciais das palavras correspondentes aos valores das anotações em inglês (e.g. ‘rb’ refere-se a ‘raised brows’). É importante manter tais convenções nesse processo de informatização do corpus, no entanto, em vista da importância de uma futura integração do corpus numa base de dados internacional.

Na verdade, a dificuldade de representação de dados de LSs não é tão distinta da dificuldade de representação das LOs. O mesmo sistema de escrita que, de um lado, facilita a transcrição das LOs, de outro lado apaga dimensões próprias da corporalidade,

como a prosódia e a gestualidade, que, se seriamente consideradas, trariam problemas similares de transcrição e produção de relatórios para os pesquisadores de LOs. A importância crescente que essas dimensões de uso da língua têm adquirido nos últimos anos, porém, tem sido acompanhada – e, em certa medida, é uma consequência – do surgimento de tecnologias que viabilizam o registro da língua e a disponibilização de gravações em seus aspectos auditivos e visuais. Como afirmam Armstrong et al. (1995: 13-14):

Looking back, it appears that linguistics was made possible by the invention of writing. Looking ahead, it appears that a science of language and communication, both optic and acoustic, will be enabled ... not by refinements in notational systems, but by increasing sophistication in techniques of recording, analyzing, and manipulating visible and auditory events electronically.

Confirmando essa previsão, a presente tese opta por utilizar os diversos recursos tecnológicos atualmente disponíveis para a representação dos dados nas análises. Ainda que se apóie sobre o sistema de notação do grupo ECS, o trabalho envolverá vários outros recursos, tais como fotos de imagens dos participantes surdos e links para arquivos de vídeo contendo os trechos analisados das gravações – arquivos esses adaptados com velocidade reduzida, legendas e recursos gráficos sobrepostos às falas originais. Afinal de contas, não há por que permanecer preso à bidimensionalidade da escrita quando hoje podemos disponibilizar ao leitor, para averiguação das análises, versões digitais de relatórios, artigos e livros, contendo, em áudio e/ou vídeo, as passagens originais analisadas (McCleary, 2003).

5.6. Conclusão

O presente capítulo discutiu aspectos metodológicos relacionados ao processo de desenvolvimento do corpus para a análise da segmentação da libras em unidades gramaticais. Em geral, discussões a esse respeito em pesquisas com LOs envolvem a explicitação de quais procedimentos e convenções, dentre várias opções, foram adotados na pesquisa e o porquê dessas escolhas. Neste caso, em que o corpus envolve uma língua de sinais, o capítulo está não apenas fazendo esse trabalho de explicitação e justificativa, mas também reportando o processo, ora marcado por tentativas bem sucedidas, ora fracassadas, de *construção* desses procedimentos e convenções. O motivo é o de que pesquisas dessa natureza são raras no Brasil e as publicações

interacionais sobre o tema são escassas mesmo nos países em que esse tipo de trabalho está mais desenvolvido.¹¹⁴ Desse modo, infelizmente, ainda é comum que os pesquisadores da área se vejam forçados a elaborar os seus próprios procedimentos independente do que está sendo produzido em outros lugares.

No que tange a questões mais gerais, apontei que a presente pesquisa favoreceu a formação de um corpus com base em dados públicos e naturais, ainda que as dificuldades técnicas de gravação de situações espontâneas de uso da libras tenham exigido estratégias metodológicas que envolveram um certo grau de manipulação. Por esse motivo, as conversas entre surdos gravadas em estúdio a pedido do pesquisador podem ser mais bem definidas como dados de natureza “semi” ou “quase-espontânea”.

Ainda a respeito de questões gerais, discutiu-se também o fato de situações de uso espontâneo da língua estarem cada vez mais sendo utilizadas por lingüistas como base para as suas análises. Essa tendência, resultado principalmente das possibilidades de registro e manipulação de dados advindas das novas tecnologias, tem sido acompanhada de críticas à adoção de metodologias que *priorizem* dados manipulados, como a eliciação de sentenças e o julgamento de gramaticalidade. No que tange ao estudo de uma língua sem escrita sobre a qual temos pouco conhecimento, como é o caso da libras, foi argumentado que o uso de dados manipulados pode comprometer diretamente a adequação dos dados obtidos e, conseqüentemente, as análises elaboradas com base nesses dados.

O processo de formação do corpus passou principalmente por três etapas: a gravação das conversas; a manipulação dos vídeos; e a transcrição dos dados. De modo geral, a primeira etapa se mostrou bem-sucedida, considerando-se a novidade da experiência e das dificuldades técnicas envolvidas no registro de uma conversa sinalizada. Apesar disso, a disposição das câmeras e a iluminação podem ainda ser melhoradas, e os problemas aqui apontados podem evitar que futuras pesquisas com enfoque similar incorram em erros similares. A etapa de manipulação dos dados envolveu não apenas procedimentos de adequação das gravações ao programa ELAN, utilizado para transcrição, mas também a exploração de recursos de vídeo que possibilitassem contornar falhas da gravação ou mesmo dificuldades de visualização

¹¹⁴ Dentre as poucas referências existentes, duas importantes são os trabalhos relativos a um *workshop* organizado pelo projeto InterSign, disponível em <http://www.sign-lang.uni-hamburg.de/intersign/> e o volume especial da *Sign Language & Linguistics*, organizado por Bergman et al. (2001).

inerentes ao processo de transcrição. Por fim, a etapa de transcrição se revelou a mais trabalhosa de todas, por envolver a anotação de uma ampla gama de trilhas em meio a um processo inevitável de adaptação ao ELAN e de elaboração do sistema de notação. Apesar disso, foi ressaltado o fato de esse trabalho intenso de observação e transcrição dos dados ter contribuído tanto para o levantamento de hipóteses de análise quanto para o desenvolvimento dos procedimentos e convenções de notação necessários ao trabalho com corpus da libras.

Por fim, foi introduzido o sistema de notação que será utilizado como ponto de referência nos capítulos de análise dos dados. A complexidade que envolve a representação de uma língua de natureza gestual-visual como a libras, contudo, demanda que o trabalho não fique preso à bidimensionalidade da escrita, de modo que, sempre que relevante, os dados da análise serão acompanhados de recursos diversos tais como: fotos da tela de transcrição do ELAN, fotos de partes relevantes dos vídeos e links para arquivos de vídeos com os trechos analisados devidamente adaptados para a demonstração das análises propostas. A idéia é a de explorar ao máximo as tecnologias existentes a fim de tornar mais claro o processo de verificação das análises aqui apresentadas.

6. Análise I: A troca de turnos na libras

6.1. Introdução

Ainda hoje, no campo das LSs, são raros os estudos de fenômenos conversacionais clássicos, como a troca de turnos e o reparo, seja com base na conversação, seja com base em outros tipos de fala-em-interação sinalizada (Baker, 1977; Coates e Sutton-Spence, 2001; Dively, 1998; McIlvenny, 1991; 1995; Van Herreweghe, 2002). Ainda mais raros são os estudos que tomam por base a conversação sinalizada em busca de conhecimento sobre a gramática das LSs (Baker e Padden, 1978; Terauchi et al., 2002). Os métodos de eliciação de sentenças descontextualizadas e apelo à intuição dos falantes têm constituído a grande maioria dos trabalhos na área, embora seja crescente também a abordagem de dados de natureza quase-espontânea, como o uso de narrativas eliciadas (e.g. Liddell, 2003a).

Parte da dificuldade de deter-se na análise da conversa em LSs é de fato tecnológica, como o capítulo sobre metodologia desta pesquisa pôde demonstrar. Sem os recursos adequados para gravação e transcrição do corpus é complicado pensar-se em qualquer estudo sério voltado para a conversação, seja com enfoque mais interacional, seja com enfoque mais gramatical. Por outro lado, especialmente de três décadas para cá, o avanço tecnológico tem diminuído significativamente essa dificuldade, de modo que a persistência em metodologias clássicas de análise revela que a opção dos pesquisadores por dados não-naturais tem dependido muito mais de posições teóricas do que de possíveis barreiras metodológicas para a formação de um corpus dessa natureza.

O presente capítulo consiste numa análise inicial sobre o processo de troca de turnos na libras. Como argumentado nos capítulos 3 e 4, a perspectiva adotada neste estudo é a de que a conversação seja um tipo de interação primordial entre os seres humanos, orquestrada entre dois ou mais participantes de modo a permitir a realização de trabalhos interacionais – que podem ter uma função pragmática mais evidente (e.g. planejar uma caça, consertar uma ferramenta) ou menos evidente (e.g. demonstrar afeto, contar vantagem). Para que esses objetivos sejam alcançados, um tipo de habilidade particular se mostra necessária: a de coordenar a atenção social sobre uma mesma “entidade” física e/ou lingüística a cada novo momento. O objetivo deste capítulo, então, é o de trazer evidências de que, na conversação em libras e possivelmente em

outras LSs, essa coordenação também deve envolver a observação do princípio do um-de-cada-vez por parte dos interlocutores.

Uma maneira de demonstrar a normatividade desse princípio é identificando momentos em que ele seja rompido, isto é, momentos em que a atenção de ambos os interlocutores sobre uma mesma “entidade” lingüística é dificultada ou impedida por circunstâncias particulares.¹¹⁵ Uma das situações típicas da conversação em que isso acontece é a sobreposição de falas. Excetuando alguns casos – que serão discutidos na seção 6.2 a seguir –, o caráter problemático das sobreposições se revela na própria fala dos participantes, que apresenta uma série de modificações desenhadas com o fim específico de solucionar aquilo que é tido como um problema interacional.

A demonstração desse caráter problemático do mais-do-que-um-a-cada-vez se mostra particularmente relevante para a atual pesquisa por um motivo principal. É comum a crença, revelada tanto em trabalhos acadêmicos quanto em especulações de leigos, de que, na conversa sinalizada, as sobreposições de fala sejam extensas e não tragam problemas para os participantes. A razão seria a de que os surdos podem observar a fala do interlocutor ao mesmo tempo em que sinalizam, sem que essa simultaneidade produza qualquer “ruído visual”. Tendo em vista que tal proposta se choca diretamente com o que tem sido articulado aqui até o momento, deverão ser trazidas evidências empíricas de que a questão relevante para a organização do um-de-cada-vez não é de ordem estritamente perceptual (capacidade de reconhecer a fala visualmente ou auditivamente), mas também de ordem cognitiva e social (capacidade de coordenar a atenção social e de oferecer contribuições seqüencialmente implicativas para o encontro conversacional).

¹¹⁵ A proposta de investigar a ordem social por meio da observação de situações em que essa ordem é rompida foi inicialmente introduzida como metodologia de pesquisa por Garfinkel (1963; 1967), que convidava seus estudantes a verificar a normatividade de certas regras do cotidiano por meio dos chamados “experimentos de ruptura” (*breaching experiments*). Alguns exemplos de experimentos dessa natureza são: ao bater-papo com uma pessoa, manter uma proximidade física maior do que a esperada; na casa da família, agir como se fosse um visitante; entre outros.

As reações sucessivas e por vezes enérgicas das pessoas diante dessas ações de ruptura mostram que tais violações estão longe de ser consideradas “banais”. Diante delas, as pessoas investiam um grande esforço no re-estabelecimento da ordem social (i.e. o seu senso de “realidade”) e, em vários casos, nem mesmo a revelação de que “tudo não passou de uma experiência sociológica” diminuía o constrangimento e a mágoa causados pela quebra da ordem social. Além disso, não apenas as “vítimas” do experimento se mostravam constrangidas em face dessas situações, mas também os próprios experimentadores, que reportavam grande dificuldade e receio de levar tais experimentos adiante.

O capítulo se inicia com a discussão de Schegloff (2000) sobre as sobreposições de fala, a fim de revelar em que circunstâncias elas se mostram problemáticas para os participantes, bem como os recursos metodicamente empregados em línguas como o inglês para solucionar esse problema interacional. Em seguida, alguns trabalhos com base na conversação sinalizada que apresentam uma maior relevância para o presente estudo serão discutidos, destacando-se as suas contribuições e também limitações de análise.

Num segundo momento, agora voltado para a análise dos dados do corpus, o objetivo inicial será o de demonstrar a importância de certas considerações analíticas e metodológicas para uma investigação adequada do fenômeno da troca de turnos em LSs. Entre essas considerações estão o estabelecimento de critérios claramente definidos para a segmentação da cadeia gestual em turnos de fala e a exclusão de certas formas de sobreposição que convencionalmente não se mostram problemáticas aos participantes. Feitas tais considerações, a análise das demais instâncias de sobreposição de fala irá revelar os principais recursos identificados na conversa em libras para a resolução desse problema interacional, com o objetivo principal de demonstrar que o modelo do um-de-cada-vez, amplamente atestado nas conversações em LOs, também se aplica a LSs como a libras.

6.2. O caráter problemático das sobreposições de fala

Num estudo que retoma e aprofunda o trabalho de SSJ sobre a sistemática do gerenciamento de turnos na fala, Schegloff (2000) discute um aspecto específico do modelo, apenas tangencialmente abordado no estudo seminal: o “dispositivo de resolução de sobreposições”. Tal dispositivo, intimamente relacionado à proposta de que a conversação se organiza a partir do princípio do um-de-cada-vez, trata da utilização metódica de determinados recursos formais pelos participantes a fim de se solucionar rapidamente as instâncias de sobreposição em favor de um dos “competidores” pela palavra.

Para SSJ, o princípio do um-de-cada-vez é, antes de tudo, fruto de uma constatação empírica: predominantemente, é assim que as conversas cotidianas tendem a se organizar. Contudo, isso não implica que desvios dessa norma não sejam fenômenos corriqueiros da conversação. Diferentemente, a proposta é a de que, independentemente da quantidade de sobreposição que possa ser identificada numa dada

conversa, os participantes ainda assim devem se orientar ao um-de-cada-vez como modelo. A evidência dessa orientação, como argumenta Schegloff, deve ser encontrada naquilo que os próprios participantes fazem da fala-em-interação quando eles se vêem diante de desvios dessa norma organizacional (2000: 3).¹¹⁶

Antes de ver o que os participantes fazem da fala nessas situações, porém, Schegloff delimita explicitamente o fenômeno da sobreposição de fala, excluindo do escopo de análise situações de fala simultânea que os próprios participantes rotineiramente não tratam como problemáticas. Em primeiro lugar, vimos que é comum, em conversas com mais do que três participantes, que a disputa pelos papéis restritos de interlocutores (i.e. falante e receptor) resulte no cisma de uma conversa em duas simultâneas – fenômeno discutido na seção 4.4 acima. Nesse contexto, a fala simultânea de participantes em conversas separadas, não importa o quão próximas entre si, tende a não produzir os efeitos típicos das falas sobrepostas tidas como problemáticas.

Em segundo lugar, é comum que a projeção de possíveis pontos de completude no turno do falante corrente resulte em transições de fala com leves sobreposições, as chamadas “sobreposições terminais”, que se estendem pelo final do turno corrente e pelo início do turno subsequente. Esse fenômeno corriqueiro, resultado do caráter indeterminado e contingente do trabalho de projeção, também não é tratado como problemático pelos participantes por envolver, na maioria dos casos, uma sobreposição programada para cessar em poucos instantes.

Em terceiro lugar, há várias práticas comuns que, embora envolvam mais do que um interlocutor participando simultaneamente, não implicam uma *disputa* pela palavra, mas sim a ocupação conjunta do espaço do turno. Entre as práticas desse tipo descritas na literatura, estão: contribuições minimalistas tais como os *continuadores* (e.g. “hm-hum”) e os *sinais de acompanhamento* (e.g. “certo”), que servem para ratificar o estatuto de falante como pertencente ao falante-corrente, concedendo-lhe um turno mais extenso; a *antecipação preliminar*, discutida na seção 3.5; a *busca colaborativa por palavras*, quando um participante é “convidado”, por meio de recursos verbais e não-

¹¹⁶ Em outros tipos de fala-em-interação, recursos e práticas de resolução de sobreposições podem ser distintos daqueles encontrados na conversação. Em debates políticos, por exemplo, a presença de um mediador e de regras explícitas parece eximir os próprios participantes do dever de regular o princípio do um-de-cada-vez. Uma consequência dessa diferença são as ocorrências comuns de extensas sobreposições de fala, nas quais os participantes agem muitas vezes como se estivessem falando sozinhos, situação que só é resolvida, então, pela intervenção desse mediador “externo”.

verbais diversos, a fornecer uma palavra que falta ao falante-corrente; entre outras atividades.

Em último lugar, certas práticas na conversação são especificamente desenhadas para serem feitas em conjunto – algumas até em unissonância –, como é o caso, por exemplo, das risadas. Como afirma Schegloff, em contextos desse tipo, ao invés de um dispositivo para resolução de sobreposição, o que se faz necessário é o conhecimento dos participantes de práticas ordenadas que garantam uma produção simultânea adequada.¹¹⁷

Excetuando-se essas condições particulares, então, instâncias de sobreposição de fala são tomadas como problemáticas pelos participantes, o que se revela pela observação de *disfluências* e *perturbações* sistemáticas quando mais do que um fala ao mesmo tempo. Schegloff utiliza o termo “disfluências” para se referir aos problemas manifestos na *progressividade da fala*, e o termo “perturbações” para os *desvios marcados da prosódia* empregada até então.

Assim, no que diz respeito à progressividade das falas em sobreposição, essas podem ser alteradas por: i) um corte abrupto, em geral por meio de uma oclusão glotal, labial ou dental brusca; ii) um alongamento ou prolongamento marcado de um som; ou iii) uma repetição de um elemento já mencionado; e quanto à prosódia das falas em sobreposição, essas podem ser alteradas por: iv) um aumento de volume; v) um aumento de pitch; ou vi) uma diminuição ou aumento repentino da velocidade da fala

¹¹⁷ Retomando um pouco a discussão da seção 3.5 sobre a função dos turnos na fala (i.e. a necessidade de se mostrarem seqüencialmente implicativos), podemos entender por que tais tipos de sobreposição rotineiramente não são tidos como problemáticos. Em conversas que sofreram um cisma, a implicação seqüencial das falas simultâneas confere a responsabilidade pelas respostas a interlocutores distintos, cada um atendendo separadamente a uma única fala. As sobreposições terminais, por sua vez, ocorrem em parte porque, em seu trabalho de projeção, os falantes são capazes de antever os términos possíveis das unidades, de modo que, na maioria das vezes, a necessidade de orientação à parte final da unidade é aliviada – a ponto de se tornar irrelevante (Auer, 2005). E quanto às contribuições minimalistas do tipo continuadores, essas têm a sua implicação seqüencial prospectiva bastante restrita, não exigindo, elas mesmas, outra resposta que não seja a própria continuação do turno em curso. Desse modo, as circunstâncias acima descritas são *rotineiramente* consideradas não-problemáticas pelos participantes pelo fato de não envolver uma forte implicação seqüencial e, desse modo, aliviar o nível atencional necessário para se responder apropriadamente a elas.

até então. Esses vários recursos podem aparecer isoladamente numa dada instância de sobreposição, ou também em combinação.¹¹⁸

Outro aspecto que Schegloff demonstra ser relevante para essa análise são as diferentes fases que caracterizam o fenômeno da sobreposição. A fase de *pré-inicialização* ocorre quando o falante corrente percebe, em meio ao seu turno em curso, por meio de pistas comportamentais do seu interlocutor, que esse está a ponto de iniciar um turno, o que deverá acarretar uma sobreposição. A *pós-inicialização* envolve o momento em que as disfluências e perturbações de um ou mais participantes demonstram o reconhecimento mútuo de que mais do que um está falando ao mesmo tempo. A *pré-resolução* ocorre quando, ao monitorar o curso da fala sobreposta de seu interlocutor, um falante se dá conta de que ela está alcançando um ponto de possível completude e que, portanto, a sobreposição poderá cessar a partir desse ponto projetado. Por fim, a fase de *pós-resolução* envolve os momentos imediatamente subseqüentes ao abandono do turno por um dos participantes, quando o seu interlocutor se vê subitamente falando em claro. Schegloff traz evidências de que, de maneira geral, os recursos formais acima listados estão diferentemente relacionados a uma ou mais dessas quatro fases, o que fortalece a hipótese de que esses sejam recursos metódicos empregados pelos participantes para resolução de um problema interacional – e não apenas reflexos de uma sobrecarga cognitiva, embora essas duas análises não sejam necessariamente incompatíveis (2000: 42).

Por fim, o último aspecto central na análise de Schegloff envolve o que ele chama de “lógica” de aplicação desses recursos formais ao longo da trajetória das falas sobrepostas. Para o autor, a grande maioria das sobreposições apresenta uma trajetória bastante breve, sendo rapidamente resolvida. O autor especifica esse tempo de resolução de sobreposições em termos do número de *batidas* (grosso modo equivalentes às unidades silábicas) necessárias até que uma sobreposição chegue ao final. A conclusão é a de que grande parte das sobreposições seja resolvida após uma única batida, isto é, tão logo se tenha evidência de que mais do que um fala ao mesmo tempo; e que, das sobreposições que sobrevivem a essa primeira batida, grande parte seja resolvida tão logo um dos participantes modifique a sua fala de modo a torná-la competitiva (i.e.

¹¹⁸ Embora Schegloff se refira apenas às *perturbações* como de natureza prosódica, aspectos relativos às ditas *disfluências*, tais como cortes abruptos e prolongamentos de sons, também podem ser interpretados como recursos dessa natureza.

empregando alguns dos recursos acima mencionados). Assim, a maioria absoluta das sobreposições já está resolvida na terceira batida de fala após a sua inicialização, o que torna as disputas mais intensas pela palavra, na verdade, casos isolados da conversação.

A análise de Schegloff sobre o dispositivo de resolução de sobreposições vem complementar a análise de SSJ sobre a organização da tomada de turnos da conversação, oferecendo uma resposta para problemas intrínsecos ao funcionamento do sistema. Como mostram as regras 1(b) e 1(c) – trazidas na seção 3.3 –, os casos de auto-seleção, em particular, abrem espaço para que, diante da possível completude de um turno em curso, mais do que uma pessoa se auto-selecione para falar. O argumento de SSJ é o de que o “primeiro a falar” deva adquirir o direito pelo turno, mas é preciso considerar que a ordem temporal das participações pode perder a sua relevância, por exemplo, quando dois ou mais participantes iniciam a sua fala juntos, ou quando essa ordem for deliberadamente ignorada por um dos competidores pela palavra, entre outras circunstâncias. O dispositivo de resolução de sobreposições, então, descreve os procedimentos convencionais utilizados pelos participantes para solucionar essas instâncias problemáticas de sobreposição, sejam elas estruturalmente previstas pelo sistema, sejam elas motivadas por razões extrínsecas a ele.

6.3. Estudos sobre conversação em línguas de sinais

Um estudo pioneiro sobre o gerenciamento de turnos na conversação em LSs foi conduzido por Baker (1977) com a ASL. Utilizando como base teórica uma combinação da proposta de Duncan (1973) e de Wiener e Devoe (1974), ambas apoiadas sobre conversação em LOs, Baker faz uma descrição minuciosa das marcas formais de caráter regulatório na ASL, isto é, os *sinais*¹¹⁹ que os falantes empregam na conversação para realizar a troca de turnos de maneira fluida. Esses sinais regulatórios envolveriam dois tipos, basicamente: aqueles que sinalizam o desejo de *manutenção* do turno corrente, empregados pelo falante e pelo receptor; e aqueles que sinalizam o desejo de *troca* de turnos, que também podem ser empregados por ambos os participantes.¹²⁰

¹¹⁹ Esses “sinais” regulatórios não devem ser vistos como, necessariamente, sinais da ASL, tal como o termo costuma ser empregado em discussões sobre LSs. Trata-se de “sinais” no sentido amplo do termo, como por exemplo a inclinação progressiva do corpo do então-receptor à frente, que, para Baker, é um “sinal” de que um falante deseja tomar um turno.

¹²⁰ Embora seja distinta de SSJ, a abordagem de Duncan não me parece antagônica a ela, mas sim complementar. O trabalho de Duncan apóia-se fundamentalmente sobre a idéia do monitoramento de

Dentre os vários sinais regulatórios analisados por Baker, cabe aqui destacar o papel que a autora atribui ao contato visual e à sua ausência. Ela afirma que a conversação sinalizada difere em particular da conversação oral pelo fato de o falante não “poder” iniciar um turno até que tenha conquistado o contato visual de seu interlocutor. Assim, vários sinais regulatórios, tais como o toque no interlocutor, gestos de apontamento e aceno com as mãos, dentre outros designados à captação da atenção nas conversas sinalizadas, seriam conseqüência direta dessa necessidade fundamental de contato visual na interação entre surdos (p. 221).

Outra questão sugerida por Baker é a idéia de que o canal gestual-visual da língua de sinais permitiria aos participantes numa conversa a produção e a compreensão da fala simultaneamente, resultando num maior grau de sobreposição do que se observa nas conversas em LOs. Como coloca Baker, “a modalidade visual da língua parece permitir aos interagentes sinalizarem e observarem os sinais do outro sem uma perda de compreensão, ao passo que nas LOs é mais difícil ouvir a fala do outro enquanto falamos” (p. 216). Nesse sentido, ela argumenta que a extensão das sobreposições nas conversas em língua de sinais parece ser maior em pelo menos um segundo em relação à extensão das sobreposições nas conversas em língua oral.¹²¹

McIlvenny (1991, 1995) é o primeiro autor a trabalhar com conversação em língua de sinais sob um olhar teórico próprio da AC.¹²² Em seu primeiro trabalho, ele se concentra em algumas considerações metodológicas sobre os desafios de estudar uma língua de modalidade visual-gestual sob o olhar da AC (i.e. dificuldades de registro de dados espontâneos, de qualidade de gravação, de transcrição dos dados, entre outras).

vários “sinais comportamentais” durante a interação, o que viabilizaria as transições fluidas entre turnos; já SSJ apóiam-se fundamentalmente sobre a idéia da projetabilidade local e contingencial do final dos turnos com base nas características de suas unidades de construção. A projeção e o monitoramento da fala são, contudo, processos que co-ocorrem no gerenciamento da troca de turnos na conversação, e, em maior ou menor grau, ambos estão presentes nas abordagens de Duncan e de SSJ.

¹²¹ Para essa conclusão, Baker se baseia no estudo de Jaffe e Feldstein, *Rhythms of dialogue*, que teria apresentado a duração média de sobreposições em LOs em torno de 0,5 segundos, em contraste com os 1,5 segundos que a autora identificou na ASL (p. 216).

¹²² Cabe fazer uma ressalva, uma vez que, em alguns momentos, tenho dúvidas quanto ao entendimento do trabalho de SSJ por McIlvenny, como, por exemplo, quando o autor afirma: “...Eu olharei para o ponto de transição dos turnos, que é o espaço no qual um sinalizador pára e outro pega a palavra” (p. 131). Essa frase é ambígua e pode implicar a interpretação (equivocada e bastante comum) de que a sistemática da troca de turnos não é feita com base na projetabilidade de *possíveis* pontos de relevância para transição, mas sim na observação empírica de que alguém parou e que o espaço está aberto para a tomada de um turno (para interpretação semelhante, ver Searle, 1992, cujo argumento foi rebatido por Schegloff, 1992).

Em seu segundo trabalho, então, ele parte para a análise propriamente dita do processo de gerenciamento de turnos entre grupos de surdos finlandeses, filmados em situações de interação espontâneas numa associação.

Diferentemente de Baker e de outros pesquisadores, McIlvenny afirma que a idéia de que as sobreposições contínuas de fala sejam permitidas na conversa sinalizada simplesmente não corresponde à observação casual de conversas na língua de sinais finlandesa (1995: 138) – embora ele não traga uma análise específica para corroborar essa observação. Por outro lado, estudando interações envolvendo três ou mais interlocutores surdos, o autor estende a proposta de Baker sobre a importância do olhar na conversa sinalizada, propondo que o acesso *local e restrito*, através do olhar, às atividades dos demais participantes na conversa sinalizada geraria práticas de gerenciamento de turnos distintas daquelas observadas em LOs.

Segundo o autor, a restrição do campo perceptual da visão, diferentemente do que ocorre com a audição, implicaria uma “exclusividade mútua entre regiões focais”, de modo que, quando um dado campo visual é focado por um surdo num dado momento da interação, outros campos visuais acabam automaticamente excluídos desse foco. Os efeitos dessa restrição seriam mais evidentes em conversas com vários interlocutores, em que a possibilidade de mais de um deles se auto-selecionar para tomar a palavra aumenta. Enquanto SSJ procuram dar conta dessas situações com a postulação de que “o primeiro a iniciar” adquire direitos pelo turno – um tema mais tarde aprofundado no trabalho de Schegloff (2000) – McIlvenny argumenta que essa regra pressupõe uma igualdade de acesso perceptual dos participantes às tentativas de auto-seleção dos demais, algo que a restrição visual inerente à interação sinalizada impediria.

Com base nessa reflexão, McIlvenny analisa duas práticas sociais que emergiriam como efeito direto dessa restrição: o *reflexo social*, i.e., a apreensão, por meio do monitoramento da reação de terceiros, de uma contribuição interacional relevante que não pode ser diretamente observada; e a *cascata*, i.e., a prática por meio da qual os vários participantes numa conversa sinalizada compartilham a responsabilidade pela chamada de atenção uns dos outros, de modo que, quando um foco de atenção unitário se faz necessário, cada surdo direciona a atenção dos participantes para os quais ele está imediatamente acessível visualmente, até que todos estejam atendendo a esse mesmo foco (p. 133-35). Embora o autor trate desses dois fenômenos em termos de práticas *conversacionais*, os exemplos analisados sugerem que

a cascata estaria mais relacionada a tipos de fala-em-interação que tendem a envolver um número maior de pessoas, como reuniões e palestras.

Um último estudo de relevância para o presente trabalho é o de Coates e Sutton-Spence (2001), que partem do pressuposto – assumido a partir da proposta de Edelsky (1981) – de que, além do modelo do um-de-cada-vez proposto por SSJ para a conversação, chamado de modelo “competitivo”, existiria a possibilidade de os falantes se orientarem a um outro modelo de organização da conversação, denominado “modelo colaborativo”.¹²³ Nesse modelo alternativo, a “palavra” não seria *disputada* mas sim *compartilhada* pelos participantes, o que tornaria a sobreposição de falas uma prática corriqueira e não-problemática.

Assim como Edelsky faz com o estudo do inglês, Coates e Sutton-Spence criticam os estudos sobre fala-em-interação em LSs, que, segundo elas, assumem de maneira pouco crítica a proposta de SSJ e adotam metodologias que favorecem a corroboração do modelo “competitivo” de organização da conversação. Elas citam o fato de Baker (1977), por exemplo, ter se apoiado sobre “falas formais”,¹²⁴ dentro das quais as várias estratégias de captação da atenção analisadas (acenos, apontamentos, toques) teriam um papel muito mais fundamental do que na conversação espontânea, na qual o uso dessas mesmas estratégias tornaria a interação “lenta e trabalhosa”, implicando uma “perda da espontaneidade” (p. 513).

¹²³ As autoras afirmam que “espaço colaborativo” (*collaborative floor*) é um termo técnico que como tal deve ser entendido, sem as conotações valorativas que o contraste “colaboração” vs. “competição” poderia invocar. Contudo, é difícil consentir com esse argumento de objetividade, uma vez que a terminologia cunhada negligencia a enorme ênfase nos estudos da AC em mostrar o caráter colaborativo de todo o processo de gerenciamento de fala na conversação.

Coates e Sutton-Spence afirmam que a terminologia por elas adotada simplesmente “tenta captar o fato de o espaço de interação colaborativo ter que ser colaborativamente desenvolvido por todos os participantes, enquanto o espaço um-de-cada-vez é desenvolvido isoladamente por um único falante” (p. 512), o que revela o entendimento profundamente equivocado das autoras sobre o trabalho de SSJ, cuja análise caminha no sentido de demonstrar que a troca de turnos fluida, com minimização de lacunas e sobreposições, é alcançada local e *interacionalmente* (SSJ: 48-51). A meu ver, problemas na interpretação de SSJ, tal como o acima citado, recorrem por todo o trabalho de Coates e Sutton-Spence.

¹²⁴ Em Baker (1977), não há qualquer referência ao fato de a conversa utilizada como base para a análise ter sido “formal”. Apenas se afirma que duas duplas de surdos foram separadamente filmadas e que, num dos casos, um dos participantes havia sido incumbido de questionar seu interlocutor sobre como se joga basquete; e, no segundo caso, um participante deveria questionar o outro sobre algo que fosse de seu interesse (p. 217). Embora tal situação não seja uma situação típica de espontaneidade, tampouco me parece uma situação de formalidade típica, aproximando-se daquilo que eu aqui tenho me referido como “dados quase-espontâneos”.

Assim, com base no pressuposto da co-existência de modelos alternativos para a organização da conversação, as autoras analisam conversas na língua de sinais britânica entre dois grupos de surdos, um masculino formado por quatro homens, e um feminino com quatro mulheres. A distribuição espacial das quatro pessoas em cada grupo foi feita de modo a dificultar que um participante pudesse visualizar facilmente todos os demais ao mesmo tempo.¹²⁵ Embora o arranjo fosse, em certo sentido, imposto pela restrição do equipamento de gravação, as autoras revelam o seu interesse particular em observar os efeitos dessa disposição espacial, tendo em vista o argumento de Baker (1977), para quem o contato visual seria crucial para a cessão, a tomada, a manutenção e o abandono de turnos em conversas sinalizadas.

A análise das autoras passa por três excertos. No primeiro, parte da conversa do grupo de homens revelaria a orientação ao um-de-cada-vez, enquanto nos dois outros excertos, um do grupo feminino e outro novamente do masculino, a conversa revelaria os participantes exercendo a palavra colaborativamente, todos participando ao mesmo tempo.¹²⁶ Apoiando-se na análise sobre o uso apurado da visão periférica por surdos (Siple, 1978; Swisher et al., 1989), as autoras argumentam que, além do uso restrito de sinais de captação da atenção prévios à tomada de turno (como os discutidos em Baker), os surdos muitas vezes iniciavam a sua fala mesmo sem garantia do contato visual dos colegas. As estratégias interacionais observadas, nesse caso, seriam a repetição contínua do movimento do sinal e a suspensão do sinal no ar até que o olhar do receptor fosse obtido (p. 523). Na interpretação das autoras, porém, tais estratégias não estariam relacionadas à disputa pelo turno, mas sim à necessidade de os sinais realizados na periferia da visão serem mais redundantes, facilitando a sua percepção pelos demais participantes (Siple, 1978). Por fim, as autoras afirmam que, ocasionalmente, pôde-se observar contribuições de um sinalizador para as quais ninguém atendia, o que revelaria de maneira ainda mais clara a orientação ao modelo colaborativo, na medida em que, dentro desse modelo, fazer parte de um espaço de interação (i.e. compartilhar um senso de coletividade) poderia ser mais importante do que trazer uma contribuição específica à conversação (p. 523, 526).

¹²⁵ Para isso, os quatro participantes foram divididos em duas duplas, cada uma sentada de um lado de uma mesa. Desse modo, um membro de uma dupla visualizava facilmente a outra dupla, mas tinha dificuldade de visualizar a pessoa sentada ao seu lado.

¹²⁶ Apesar disso, segundo as autoras, os homens demonstrariam “preferência” pelo modelo do um-de-cada-vez, ao passo que as mulheres demonstrariam “preferência” pelo modelo “colaborativo”.

Se, de um lado, os estudos acima resumidos certamente contribuem com reflexões e análises interessantes para o estudo da conversação em LSs, de outro lado as diferentes análises apresentam alguns pontos fracos, ora em sua articulação teórica, ora na elaboração metodológica, ora na própria interpretação dos dados. Considerando esses aspectos, pretendo me concentrar na crítica de uma questão que percorre todos os estudos citados, independentemente de suas diferentes conclusões: o modo como a especificidade da conversa sinalizada é analisada em termos basicamente perceptuais, colocando-se de lado as dimensões cognitivas e interacionais inerentes ao estabelecimento do encontro conversacional.

As duas questões sobre a relação entre visão e audição, para as quais cada um dos trabalhos citados oferece uma resposta distinta, poderiam ser colocadas da seguinte maneira: i) a sobreposição de vozes nas LSs teria uma natureza distinta da sobreposição de vozes nas LOs, tendo em vista que olhar a sinalização de outra pessoa enquanto sinalizamos parece não produzir um “ruído visual” da mesma maneira que o som de duas vozes simultâneas produz um “ruído sonoro”?; ii) as limitações perceptuais do foco visual teriam alguma implicação diferenciada para a organização da tomada de turnos nas LSs, já que o fato de uma fala oral poder ser percebida por um participante independe da direção de onde ela venha, ao passo que a percepção da fala sinalizada se mostra restrita a um campo visual mais limitado?

A resposta de Baker (1977) para ambas as questões é “sim”. Para a autora, o fato de o surdo ser capaz de observar o seu interlocutor enquanto ele próprio sinaliza faz com que a conversação sinalizada apresente sobreposições de fala maiores; além disso, o fato de a visão ter um campo perceptual mais restrito faria do contato visual um sinal fundamental para o gerenciamento da troca de turnos. McIlvenny (1995), por sua vez, diz “não” à primeira questão, afirmando que o princípio do um-de-cada-vez pode ser observado também na conversa sinalizada; e “sim” à segunda questão, estendendo a análise de Baker por meio da descrição de práticas sociais (i.e. o reflexo social e a cascata) que seriam decorrentes da “mútua exclusividade entre focos de visualização” na conversa sinalizada. Por fim, Coates e Sutton-Spence (2001) propõem que as análises de conversas sinalizadas têm desconsiderado a acuidade da visão periférica dos surdos, muito mais apurada do que a dos ouvintes, o que livraria os interlocutores surdos da necessidade de estabelecer um contato visual direto para organizar a sua participação na conversação e favoreceria a participação de *todos-a-cada-momento*.

Em todos esses casos, ainda que eventualmente para alcançar conclusões distintas, nota-se que a discussão sobre a especificidade da conversa sinalizada gira principalmente em torno da possibilidade ou impossibilidade de *acesso perceptual* dos surdos ao que ocorre numa conversação, em detrimento de todo o trabalho cognitivo e interacional – discutido nos capítulos 3 e 4 – que necessariamente está relacionado ao processo de gerenciamento de turnos e de construção de sentidos na conversação. Como tem sido argumentado neste trabalho, a compreensão, seja de enunciados visuais numa língua de sinais, seja de enunciados auditivos numa língua oral, envolve muito mais do que o acesso e o reconhecimento perceptual desses enunciados. Crucialmente, é necessário que o nosso foco de atenção esteja direcionado a uma dada fala para que possamos, de um lado, fazer sentido dela dentro da seqüência interacional imediata, e de outro, oferecer uma resposta apropriada dentro das restrições de participação inerentes à conversação.

No caso de Baker (1977), seus próprios achados parecem contradizer a idéia de que os surdos admitam uma sobreposição de fala contínua na conversação. A autora faz uma descrição detalhada de inúmeros sinais “regulatórios”, desde disfluências sutis da fala sinalizada até pedidos explícitos de atenção, que são empregados pelos participantes surdos para gerenciar a sua participação na interação. No que diz respeito à análise de que o período médio de sobreposição de fala seria maior na conversa sinalizada, então, tal reflexão deve ser reconsiderada em vista de algumas considerações metodológicas e teóricas. Do ponto de vista metodológico, seria fundamental a qualquer análise sobre sobreposição em LSs trazer uma definição clara de quando a fala sinalizada se inicia e quando ela é encerrada, tendo em vista que, nas LOs, essa questão é sempre pressuposta considerando-se os períodos audíveis de vocalização – ver discussão na seção 2.5 sobre pausas nas LSs.¹²⁷ Do ponto de vista teórico, seria fundamental que a análise quantitativa sobre as extensões de sobreposições fosse acompanhada de uma análise qualitativa dos dados. Tendo em vista que o princípio do um-de-cada-vez é observado pelos falantes mesmo quando os turnos são de fato sobrepostos – o que se manifesta em várias disfluências e perturbações nas falas –, a

¹²⁷ Nas LOs, aspectos gestuais não acompanhados de vocalização verbal – tal como inspirações de ar audíveis que antecedem a tomada de turnos (Schegloff, 1996) ou a mudança postural e a gesticulação (Schegloff, 2000) – são considerados nas *imediações* das sobreposições (e.g. fase de pré-inicialização), mas não como capazes de constituir sobreposições em si (ver discussão na seção 6.4.1 abaixo).

análise qualitativa se mostra imprescindível para qualquer afirmação sobre diferenças na forma de organização da conversação sinalizada. Parte dessa análise qualitativa envolveria, entre outras coisas, uma reflexão sobre o estatuto dos diferentes tipos de sobreposição identificados na conversa, uma vez que nem todos se mostram problemáticos aos participantes – ver discussão da seção 6.2 acima.¹²⁸ Desse modo, a análise de Baker não pode ser tomada como evidência teórica e metodologicamente robusta de que o um-de-cada-vez não se aplique à conversa sinalizada, tampouco de que os falantes de LSs sejam capazes de produzir e compreender língua ao mesmo tempo.

McIlvenny (1995), por sua vez, aponta os fenômenos do reflexo social e da cascata como efeitos daquilo que ele considera uma restrição específica da conversação sinalizada em contextos de auto-seleção de falantes: a exclusividade-mútua de focos de visualização. Mas será que esses dois fenômenos são, de fato, exclusivos da interação sinalizada? Um olhar casual sobre situações de conversação entre ouvintes revela, por exemplo, que certas mudanças comportamentais salientes por parte de um dos participantes, tal como uma mudança de olhar repentina, pode acarretar um reflexo social nos demais participantes, que, em vista do monitoramento contínuo de seu interlocutor, mudam também o olhar para a mesma direção.¹²⁹ Igualmente, em situações de interação que envolvam muitos participantes e uma necessidade de centralização do foco de atenção (e.g. reuniões, aulas, palestras), é comum observarmos a distribuição do trabalho de captação de atenção entre ouvintes, partindo de um participante central para os demais como numa reação em cadeia, até que todos estejam voltados para o mesmo foco (e.g. os pedidos de silêncio em salas de aula, que em geral partem do professor mas que costumam ser feitos também por alguns alunos quando a dispersão é grande).

¹²⁸ Baker chega a evocar a distinção de Duncan (1973) entre “turnos simultâneos” (quando os dois participantes ambicionam a palavra) e “fala simultânea” (quando um participante produz sinais de acompanhamento diante de um turno que está em curso) (p. 216). Porém, a análise sobre a extensão das sobreposições acaba não sendo relacionada a um ou outro tipo de sobreposição especificamente, e outras formas de sobreposição não problemáticas – tais como as discutidas na seção 6.2 – não são abordadas.

¹²⁹ Aqui, faço uma breve digressão para apontar um viés muito comum nas análises contrastivas entre surdos e ouvintes no que diz respeito a diferenças linguísticas, culturais, cognitivas e sociais. Em geral, contrapõe-se fortemente o caráter “visual” dos surdos ao caráter “auditivo” dos ouvintes, esquecendo-se o fato de que a maioria dos ouvintes não apenas é capaz de ver como também se apóia na visão como um dos seus sentidos fundamentais para a construção do conhecimento. No contra-exemplo aqui apresentado, fica claro que o monitoramento do outro na conversação entre ouvintes não envolve apenas sinais comportamentais auditivos, mas também visuais (ver discussão da seção 4.5 sobre o papel da gestualidade no gerenciamento da atenção conjunta).

Diante dessas observações, nota-se que os efeitos da diferença entre surdos e ouvintes no gerenciamento da troca de turnos não é uma diferença qualitativa, mas de grau. Na medida em que, quando comparada à audição, a visão envolva um campo perceptual menor *numa dada interação*, a ocorrência de fenômenos tais como o reflexo social e a cascata será maior numa conversa entre surdos do que numa conversa entre ouvintes. Desse modo, ainda que *típicas*, tais práticas certamente não devem ser vistas como *específicas* da conversação sinalizada.¹³⁰

O argumento de McIlvenny é o de que, no processo de auto-seleção de um participante, a acessibilidade restrita dos surdos aos turnos emergentes (e possivelmente conflitantes) em determinados TRPs, assumida pelo modelo de SSJ, não poderia ser pressuposta no caso da conversação sinalizada. De fato, tal acessibilidade parece ser assumida no modelo quando os autores descrevem, com relação à regra 1(b), que o “primeiro a iniciar [a auto-seleção] adquire direito ao turno”. Contudo, isso não impede que a conversação oral também esteja sujeita a situações em que aquele que conquista o turno por meio de auto-seleção não seja objetivamente “o primeiro a iniciar”, mas sim alguém que simplesmente não ouviu o início do turno de outro participante e acabou vencendo a disputa pela palavra.¹³¹

¹³⁰ Quando enfatizo as particularidades de “uma dada interação”, faço-o porque as circunstâncias imediatas podem beneficiar, ora surdos, ora ouvintes, no que se refere à abrangência de seu campo perceptual, dependendo de cada situação. Por exemplo, a visão parece ser limitada a um ângulo de aproximadamente 180° em relação aos nossos olhos, ao passo que a audição teoricamente não se restringe a uma circunferência limitada de acesso ao som; por outro lado, a visão parece ter um potencial maior de alcance para a fala produzida a distância do que a audição. Além disso, diversos fatores ambientais (e.g. o nível de ruído visual ou auditivo no ambiente) ou fisiológicos (e.g. problemas visuais/auditivos temporários ou permanentes dos participantes) podem afetar a acessibilidade aos turnos de outros falantes de maneiras que não são específicas de uma modalidade ou outra de língua. Por fim, cabe assinalar que tanto surdos quanto ouvintes buscam sempre se adaptar às situações sociais de interação a fim de expandir ao máximo as suas condições de acessibilidade das falas dos demais participantes (e.g. formando pequenos círculos em que o monitoramento mútuo é facilitado).

¹³¹ Além disso, se as análises de sobreposição de fala em LOs dessem maior consideração à gestualidade, acredito que certas formas de expressão corporal também poderiam ser analisadas em termos de “inicialização de fala sobreposta” – em geral atribuída exclusivamente ao início da vocalização. Por exemplo, em alguns casos, uma sensação de “interrupção” pode ser sentida por um participante que não chegou sequer a produzir um som, mas que produziu pistas gestuais suficientes para indicar um claro desejo de tomar o turno (e.g. inspirar ar, abrir a boca, assumir uma postura mais tensa, levantar a cabeça e a sobrancelha). Se tais sinais corporais passassem a ser considerados nas análises de sobreposição em LOs, uma constatação comum seria a de que os ouvintes estão igualmente constrangidos por um acesso perceptual mais restrito à inicialização da auto-seleção por parte de seus interlocutores. Mais uma vez, tais argumentos sugerem que a efetivação das trocas de turnos numa conversação dependa de fatores

A crítica central sobre a análise de McIlvenny, contudo, não reside nas diferenças de grau entre os campos de percepção visual e o auditivo. O ponto central é o de que, mais do que uma exclusividade mútua de foco *visual*, o que o gerenciamento de turnos na conversação implica é uma restrição mais difícil de ser contornada, de natureza cognitiva e interacional: a exclusividade-mútua de *focos de atenção social*. Nesse sentido, tanto uma fala oral quanto uma fala sinalizada podem deixar de concretizar uma participação na conversação, não porque não tenham sido *percebidas*, visual ou auditivamente, mas porque não tenham sido *atendidas* pelos participantes de maneira suficientemente centrada a fim de possibilitar uma resposta que se mostre seqüencialmente implicativa e temporalmente coordenada.¹³²

Por fim, Coates e Sutton-Spence argumentam contra a idéia, tanto de Baker quanto de McIlvenny, de que o contato visual seja necessário para a efetivação das trocas de turnos em conversas sinalizadas. Para fortalecer a sua proposta, de que os surdos – em especial as mulheres – participam orientando-se a um modelo “colaborativo” de todos-juntos-a-cada-momento, as autoras trazem como argumento adicional a maior acuidade da visão periférica em surdos, o que possibilitaria a captação da sinalização de mais de um interlocutor sem a necessidade de contato visual direto. Novamente, aqui, a crítica cabível é a de que ser capaz de *perceber* sinais na região periférica da visão não garante que as falas “periféricas” possam ser *atendidas*, *compreendidas e respondidas* de modo a se mostrarem seqüencialmente implicativas. Evidências dessa limitação atencional podem ser encontradas nos próprios dados trazidos pelas autoras, que, em várias instâncias de sobreposição, revelam disfluências e

extrínsecos ao sistema que podem ser menos ou mais favoráveis dependendo das diferentes situações de interação, e não de ela ser oral ou sinalizada.

¹³² Ao final de seu estudo, McIlvenny afirma: “Um resultado importante aqui é o de que as atividades dos receptores e o tempo das espiadas, das transições de olhar e do contato visual mútuo são sistematicamente relevantes para a resolução da transição entre sinalizadores; não cabe apenas aos sinalizadores em competição resolver uma sobreposição potencialmente problemática” (p. 143). Sob o olhar da discussão acima, parece-me que a contribuição de McIlvenny não foi tanto a de constatar uma especificidade da conversa sinalizada, mas sim a de chamar a atenção para um aspecto do modelo de SSJ que carece de aprofundamento: como é que se resolvem os conflitos por turnos nas situações em que múltiplos participantes se auto-selecionam. Em seu trabalho sobre sobreposição, Schegloff (2000) sugere diferentes formas de resolução que parecem envolver exclusivamente os falantes em competição, embora o autor mencione de passagem uma possível relevância do olhar e do corpo (p. 8). O trabalho de McIlvenny sugere que os demais participantes, a princípio “de fora da competição” pela palavra, podem assumir um papel crucial nesse processo de resolução, em especial pelo direcionamento do olhar, mas eventualmente também por outros sinais comportamentais e práticas convencionizadas.

reciclagens (e.g. repetições dos movimentos e manutenção da suspensão de sinais em sobreposição, p. 523, 527).¹³³

Além disso, a afirmação de que as falas sobrepostas dos participantes em muitos casos não envolveram quaisquer sinais de perturbação torna-se questionável em razão de alguns cuidados metodológicos que não foram tomados. Dentre esses cuidados, o principal refere-se à consideração das implicações que o número de participantes na conversação poderia acarretar. As seções de conversação gravadas pelas autoras, formadas por grupos de quatro participantes, apresentavam uma propensão contínua ao cisma das conversas. Esse potencial exigiria um olhar diferenciado para a questão da orientação ao um-de-cada-vez, bem como para os reflexos dessa orientação na fala dos participantes diante das sobreposições de turnos. Uma vez ignorado esse fator, aquilo que poderia se desenvolver como a sobreposição de falas pertencentes a duas conversas distintas (porém próximas) se mostra passível de ser interpretado como uma forma distinta de organização numa única conversa, quando os participantes exercem a palavra “colaborativamente” e “todos atendem a todos ao mesmo tempo”.¹³⁴

¹³³ É curioso ver como as autoras interpretam com tranquilidade uma passagem, a meu ver altamente disputada, como uma evidência de um modelo antagonístico ao do um-de-cada-vez (p. 516-7). Sobre a passagem em questão, Coates e Sutton-Spence afirmam: “É *óbvio* que as quatro amigas não estão organizando a sua fala de acordo com o modelo do um-de-cada-vez; se elas estivessem, elas seriam culpabilizadas pelas freqüentes violações do direito de cada uma pelo turno, e no entanto ninguém protesta e há uma ausência notável de reciclagem das falas sobrepostas que é encontrada na conversa quando o falante se sente interrompido” (p. 517, ênfase minha). Essa constatação “óbvia” se refere a uma passagem em que pude contar cinco instâncias de HEY (i.e. um aceno de mãos para chamar a atenção dos demais) realizadas por três das quatro participantes; seis instâncias de alongamentos dos sinais em sobreposição, mantidos em suspensão no ar por um período de tempo significativamente longo; e frases que, a despeito da falta de detalhamento da notação, parecem instâncias claras de interrupção no desenvolvimento de uma idéia (como na fala de Tanya: “FUNNY SEE PAM ME HEY NO HEY”, que pelo contexto imagino poder ser traduzida por algo como: “É engraçado, eu vi a Pam= Hey!... Não!... Hey!”; mas que as autoras optaram por traduzir substituindo os sinais regulatórios por reticências (“It’s funny, I saw Pam...”, p. 517). Assim, ao contrário do que a “constatação óbvia” das autoras sugere, e mesmo em face da carência de informações sobre modulações dos sinais no sistema de notação e/ou na discussão, parece-me que há nessa passagem uma clara tentativa de se solucionar a sobreposição de fala.

Além disso, Coates e Sutton-Spence parecem ter uma visão muito restrita sobre os efeitos da sobreposição de turnos nas falas dos participantes que se orientam sobre o um-de-cada-vez. Como a citação acima trazida revela, elas apontam como efeitos desse fenômeno apenas dois aspectos: a) os protestos explícitos contra a sensação de “interrupção” de um turno (que são raros nas conversações, como aponta Schegloff, 2000: 51 e 2002) e b) as reciclagens de parte da fala sobreposta (que é apenas um dos inúmeros efeitos que se manifestam em sobreposição de fala, como mostra a discussão da seção 6.2).

¹³⁴ É curioso notar que Coates e Sutton-Spence, seguindo a crítica de Edelsky (1981), argumentam contra o uso de díades em análises de conversação, sob o argumento de que análises baseadas nesse tipo de corpus favorecem o modelo do um-de-cada-vez. Esse argumento, porém, não encontra qualquer

Nesse aspecto, a observação do direcionamento do olhar, como sugerem Baker e McIlvenny, poderia ter um papel central para favorecer uma ou outra interpretação. Apesar disso, as autoras concluem o seu trabalho com a afirmação de que, nos excertos estudados, os participantes estavam “todos direcionando a atenção uns aos outros em todos os momentos, embora a cada novo momento o olhar tivesse de estar direcionado a um sinalizador e não aos demais” (p. 525). Problemas teóricos e metodológicos tais como os acima apontados, contudo, comprometem quaisquer conclusões alcançadas sobre a inoperância do modelo um-de-cada-vez.

6.4. Sobreposições de fala na libras

6.4.1. A segmentação da cadeia de fala sinalizada

Uma primeira consideração necessária para uma análise adequada de sobreposições na conversação em libras é a definição precisa de onde se inicia e onde termina a fala dos surdos propriamente dita. Como mencionado na seção 2.5, pesquisadores de LOs têm assumido que a sobreposição envolva apenas os momentos em que duas vocalizações em concomitância são identificadas. Formas de gestualidade que acompanham necessariamente essa vocalização (i.e. as articulações do trato vocal) ou que a acompanham opcionalmente (i.e. gesticulações com o braço, movimento de sobrancelha, cabeça, etc) não são vistas como, em si, passíveis de acarretar sobreposição.

Nas LSs, diferentemente, a distinção entre a “vocalização” (i.e. a parte “significativa” da fala) e os “gestos que a acompanham”, opcional ou necessariamente, não é nada nítida, uma vez que, ao contrário das LOs, a “matéria” da percepção lingüística são os próprios gestos que se observam realizados diante de si, e não o sub-

fundamentação em análises de sobreposições. Schegloff (2000), que analisa dados provenientes de conversas com diferentes números de participantes, alcança uma conclusão inteiramente distinta: “[R]evela-se, com grande regularidade, que, quando mais do que uma pessoa está falando num dado momento, duas pessoas estão falando, e não mais”; e, em seguida, que “a fala por mais do que duas pessoas parece se reduzir à de duas pessoas (ou à de uma) de maneira ainda mais eficiente do que a fala de duas é reduzida à de uma” (p. 7). Assim, não há qualquer evidência de que uma conversa com um número maior de pessoas afete a estrutura organizacional da troca de turnos (alocada sempre de maneira binária) implicando um maior número de sobreposições, a menos que se considere todos os diferentes tipos de sobreposição de maneira indiscriminada.

produto sonoro desses gestos.¹³⁵ Desse modo, tanto os gestos articulatórios que fazem a transição entre o repouso e a “fala de fato” – e entre os vários sinais numa dada fala – quanto os gestos “opcionais” que podem acompanhar os sinais são produzidos pelo mesmo canal de fala sinalizada.

Para uma análise adequada das trocas de turno e das sobreposições na conversação em libras, então, uma primeira etapa necessária para a “limpeza do terreno” é a segmentação da cadeia gestual em fases. Como visto na seção 5.4, os estudos da gestualidade mostram que a cadeia gestual pode ser segmentada, grosso modo, em fases de *preparação*, fases *expressivas* (que podem ser formadas, ou por uma *suspensão independente*; ou por um *golpe*, opcionalmente precedido e/ou sucedido de *suspensões dependentes*) e fases de *retração*. Entender a função que cada uma dessas fases desempenha na fala sinalizada, em contraste com o que se observa nas LOs, é fundamental para que uma comparação inter-modal sobre os fenômenos em questão possa ser apropriadamente realizada.

A análise da passagem apresentada na figura 20 é um ponto de partida útil para essa reflexão.¹³⁶ <Nesse excerto>, W está falando sobre sua dificuldade com uma apresentação de física, que ele teve que fazer na escola para o professor junto a um grupo de colegas. Em meio à fala de W (“Que difícil e complicado!”, trilha W.1), R inicia um turno com um questionamento – aparentemente um reparo para garantir a sua compreensão sobre o motivo do receio e da ansiedade de W:

¹³⁵ O chamado efeito McGurk mostra que, também nas LOs, não apenas o som se mostra relevante lingüisticamente, mas também os gestos articulatórios que o produzem. Bolinger (1982b), similarmente, mostra que a interação entre os gestos entoacionais, faciais e corporais é determinante na interpretação global dos enunciados. Apesar disso, é difícil negar a maior relevância da dimensão sonora nas LOs em detrimento da dimensão gestual, quando consideramos, por exemplo, o grande potencial que temos para nos comunicar sem atender visualmente ao nosso interlocutor, em contraste com a relativa limitação que temos para nos comunicar quando os gestos corporais são a única fonte de informação.

¹³⁶ À esquerda da figura, as letras W e R referem-se às iniciais dos nomes dos falantes, e os números que as acompanham são utilizados para futura referência às trilhas, podendo ser modificados de figura para figura dependendo do número de trilhas apresentado em cada uma delas. As letras maiúsculas na transcrição representam as glosas dos sinais. Nessa figura, a trilha W.2 envolve glosas referentes a sinais não-manuais (e.g. feitos com a cabeça), ao passo que as trilhas W.3 e R.2 envolvem glosas referentes a sinais manuais.

A segmentação da cadeia de fala em frases (trilha W.1 e R.1) foi feita com base numa análise das características formais das unidades prosódicas na libras. Essa segmentação será pressuposta na presente discussão e argumentos para justificá-la serão encontrados no capítulo 7.

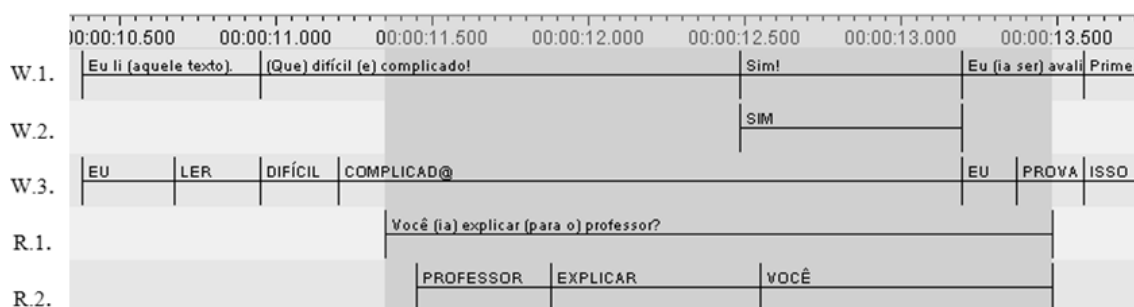


Figura 20. Ocorrência de sobreposição sem consideração das fases do gesto.

A faixa vertical escura na figura 20 indica a extensão da sobreposição neste trecho *sem qualquer* consideração sobre as diferentes fases do gesto. Ela mostra que R inicia o seu turno no momento em que W realiza o sinal COMPLICAD@, como indicado no limite esquerdo da faixa vertical escura.¹³⁷ Próximo aos 12s500ms, W responde ao questionamento de R com um “Sim”, realizado por meio de um sinal não-manual (i.e. um aceno enfático de cabeça, glosado na trilha W.2 como SIM). Próximo aos 13s500ms, então, a sobreposição cessa para ceder lugar à fala de W, que segue em claro a partir daí. Sem uma análise específica sobre as fases do gesto, isto é, conferindo-se valor indiscriminado a toda a unidade gestual produzida por R (i.e. desde o momento em que suas mãos saem do repouso até o momento em que elas retornam totalmente ao repouso), a sobreposição na passagem revela 2s135ms¹³⁸ de falas simultâneas.

Consideremos agora a mesma passagem acrescentada da anotação sobre as fases do gesto, como mostra a figura 21.¹³⁹ Não há dúvidas quanto ao fato de a fase expressiva, que carrega a carga semântica do sinal, ser relevante para o fenômeno da sobreposição, tendo em vista que a implicação seqüencial de uma dada fala pode ser apenas avaliada e apropriadamente respondida atendendo-se a esse conteúdo semântico. Mas e quanto às fases de preparação e retração das unidades gestuais e dos sinais que as compõem?

Analisando inicialmente a fase de retração, nota-se que, após R iniciar um turno que resulta numa sobreposição, W alonga o sinal COMPLICAD@ por meio de uma

¹³⁷ A lacuna entre o início da tradução, em R.1, e o início da primeira glosa da frase traduzida, em R.2, atribui-se ao fato de, nessa passagem, R iniciar a sua fala com um gesto não-manual que não aparece na transcrição (levantamento da sobrancelha), antecedendo o início do sinal manual PROFESSOR.

¹³⁸ Os valores relativos à duração das fases nesta seção foram extraídos do ELAN.

¹³⁹ As trilhas W.4 e R.3, referente às fases do gesto, envolvem as seguintes siglas de notação: “pr” (*preparation*); “st” (*stroke*); “rt” (*retraction*); “hb” (*hold before stroke*); e “ha” (*hold after stroke*).

suspensão pós-golpe (“ha”, trilha W.4, 12s050ms), seguida de uma longa retração parcial (o “rt” que inicia sob a faixa escura vertical da esquerda e se estende até os 13s200ms).¹⁴⁰ Meu argumento é o de que, nessa fase do gesto, a fala em questão não tem o mesmo estatuto de uma fala plena e, assim sendo, não deve ser contada como instância de sobreposição no sentido que vem sendo tratado neste trabalho.

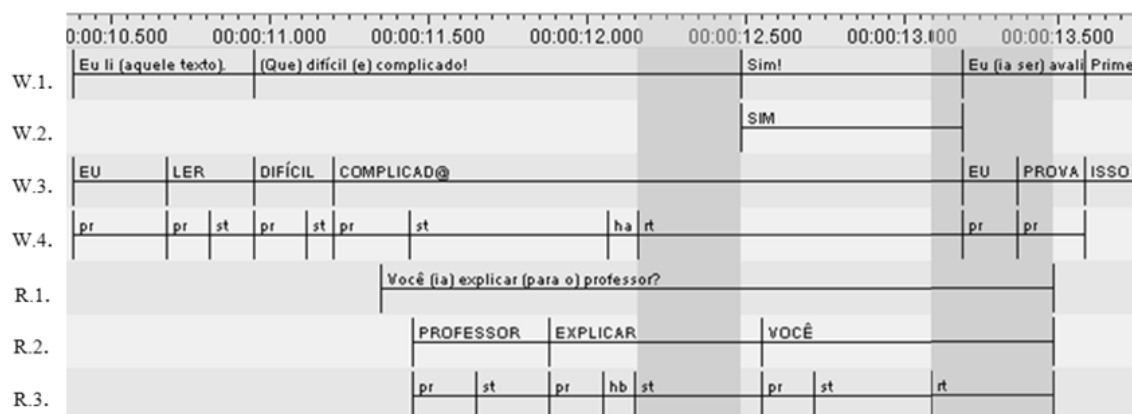


Figura 21. Ocorrência de sobreposição considerando-se a fase de retração.

Na passagem, são duas as evidências de que casos como esse – em que se observa uma retração parcial dos sinais no ar – não devam ser considerados fala no sentido pleno do termo. Em primeiro lugar, W revela uma clara orientação à R, fitando-a de maneira mais intensa à medida que o sinal COMPLICAD@ vai se tornando progressivamente relaxado. Em segundo lugar, como se revela na transcrição entre 12s500ms e 13s200ms, o sinal COMPLICAD@ permanece estático, em retração parcial, sem qualquer alteração em sua dinâmica, mesmo durante a resposta subsequente de W ao questionamento de R por meio de um sinal não-manual (SIM, trilha W.2, 12s500ms). Seria certamente equivocada pensar, sob essas circunstâncias, que o “resto” do sinal COMPLICAD@ estivesse trazendo qualquer contribuição de sentido à resposta de W.¹⁴¹

¹⁴⁰ Na retração *total*, as mãos retornam à posição de repouso – que, no caso das conversas gravadas, localiza-se no colo dos participantes. Já na retração *parcial*, observa-se um afrouxamento da configuração das mãos acompanhado de um lento deslocamento das mãos no ar, sem qualquer direcionamento específico ou vagamente direcionadas à posição de repouso.

¹⁴¹ Na verdade, é como se W tivesse “esquecido” o sinal COMPLICAD@ no ar, enquanto presta atenção à fala de R (faixa vertical escura da esquerda) e enquanto responde a esse questionamento com um “Sim” (dos 12s500ms aos 13s200ms).

A idéia de que W “esqueceu” o sinal no ar, porém, não deve ser levada às últimas conseqüências. Embora eu não considere sinais que pairam no ar como fala no sentido pleno do termo, também não me parece correto interpretá-los como pausas plenas – essas reveladas mais claramente quando as mãos do

A outra instância de retração aparece na trilha R.3, sob o sinal VOCÊ. Respondendo a R com um SIM (trilha W.2), W inicia a sua contribuição ao final do sinal EXPLICAR de R (trilha R.2), antecipando o final da unidade de maneira bem-sucedida. O resultado dessa antecipação é uma sobreposição terminal entre o sinal SIM de W e o sinal VOCÊ, produzido por R (trilha R.2). É interessante notar, então, que W dá continuidade ao seu turno (com a frase “Eu ia ser avaliado”, trilha W.1) tão logo a fala de R inicia a fase de retração total, como mostra a faixa vertical escura da direita – mais precisamente, 0,1 segundo após o início da retração do sinal VOCÊ. Essa observação nos permite constatar que o sinal SIM, de W, perdurou pelo período de tempo exato que R levou para alcançar o fim da unidade – terminada no ponto em que o sinal VOCÊ apresenta uma retração total.¹⁴²

Assim, da mesma maneira que não seria correto considerar como instâncias de sobreposição a simultaneidade entre a pergunta de R (PROFESSOR EXPLICAR VOCÊ?) e a retração parcial da fala de W (após o golpe em COMPLICAD@), também não seria correto considerar como sobreposição a simultaneidade da frase de W (EU PROVAR) e a retração total da fala de R (após o golpe em VOCÊ). Ambos os fragmentos, que aparecem sob as duas faixas verticais escuras, devem ser

falante retornam totalmente à posição de repouso.

Talvez uma explicação plausível para o fenômeno de sinais que pairam no ar seja aquela proposta por Local e Kelly (1986) para os diferentes tipos de silêncio nas LOs. Como discutido na seção 3.5, Local e Kelly apontam duas maneiras distintas de “pausar” ou “fazer silêncio” em seu estudo sobre a exploração de recursos fonéticos no gerenciamento da troca de turnos: na primeira, não há qualquer movimento articulatório do aparelho fonador durante o silêncio, o que, para os autores, funciona como uma pista de que o então-falante deseja ceder o turno; na segunda, há um movimento articulatório de fechamento glotal durante o silêncio, o que, para os autores, indica o desejo do falante-corrente de permanecer com o turno, a despeito da pausa. A evidência dessas análises estaria na mudança sistemática de turnos que se opera frente aos silêncios plenos, em contraposição à manutenção sistemática dos turnos com o falante-corrente quando ele faz uso de “silêncios glotais”.

Nas LSs, a manutenção dos sinais no ar por meio de períodos de retração parcial talvez tenha uma função similar: indicar uma pausa, porém uma pausa que ainda revela um desejo de dar continuidade ao turno de fala.

Essa situação não deve ser confundida com o recurso produtivo e significativo nas LSs de manter “pedaços de sinais” no ar enquanto a outra mão realiza outros sinais (e.g. para salientação de um elemento num contraste, ver discussão da seção 7.4.2). Nesse caso, a manutenção de “restos” de sinais no ar, diferentemente das ocorrências estudadas nesta seção, é deliberada e não apresenta as características formais que definem a fase de retração parcial (i.e. afrouxamento da configuração de mão, acompanhado de deslocamento lento e sem direção das mãos no espaço).

¹⁴² Outras evidências de que a retração já não faz parte da unidade produzida por R são <[alguns sinais não-manuais que quebram com a coesão facial e corporal da pergunta de R](#)>: na fronteira entre o final do golpe e o início da retração, uma piscada de olhos; e, no restante da retração, os breves e contínuos acenos de cabeça que vão acompanhar todo o turno subsequente de W.

apropriadamente excluídos da instância de sobreposição por envolverem uma fase do gesto que se mostra pouco ou nada implicativa sequencialmente.

No que diz respeito à fase de preparação, nota-se que a fala de R se inicia 0,3 segundos antes de o primeiro sinal (PROFESSOR) alcançar a fase expressiva, como revela a faixa vertical escura na figura 22. Esses 0,3 segundos envolvem a preparação do turno de R, inicialmente relativa aos sinais não-manuais da pergunta sim/não (i.e. levantamento da sobrancelha e o início de uma inclinação da cabeça para baixo) e, em seguida, relativa ao sinal manual PROFESSOR, como mostra a anotação “pr” na trilha R.3. Além disso, a fase de preparação ainda recorre em três outras instâncias na passagem em questão. Duas delas referem-se às preparações dos dois sinais manuais subseqüentes na pergunta de R (EXPLICAR e VOCÊ), e a última refere-se à preparação da frase com a qual W dá continuidade ao seu turno (EU PROVAR), ao término da sobreposição com R.

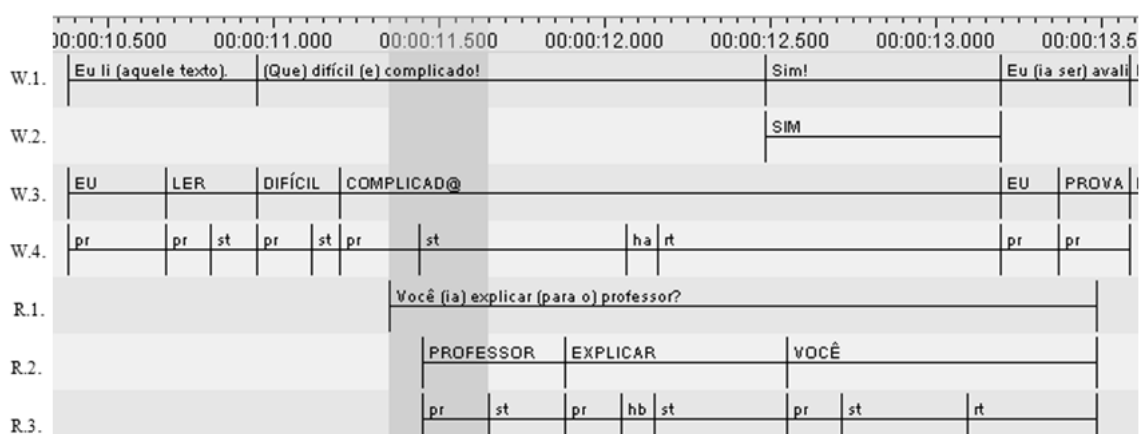


Figura 22. Ocorrência de sobreposição considerando-se a fase de preparação.

A meu ver, a despeito das semelhanças entre as fases de preparação e retração – no sentido de ambas não possuírem conteúdo semântico e de ambas se mostrarem dependentes da fase expressiva – cada uma dessas fases desempenha funções bastante distintas na interação. Na fase de retração, o foco de atenção do sinalizador já foi inteiramente mudado – para a inatividade das mãos (i.e. o papel de receptor na conversação) e, eventualmente, para a realização de um novo sinal/enunciado. Na fase de preparação, diferentemente, entendo que a atenção do sinalizador deva estar diretamente voltada para a realização de um novo sinal/enunciado, de modo que o investimento cognitivo na realização da preparação de um sinal deve comprometer o investimento cognitivo necessário para se atender a uma fala simultânea do interlocutor.

Embora a passagem em questão não traga fortes evidências para essa hipótese, o modo como W volta o olhar a R no período preparatório de 0,3 segundos que antecede o golpe do sinal PROFESSOR, e como ele permanece fitando R durante a realização dos sinais subseqüentes, sugere que, de fato, os interlocutores demonstrem orientação à fase de preparação dos sinais em pelo menos dois sentidos: quando tal fase envolve o início de uma *unidade gestual* (i.e., uma preparação que tem como ponto de partida as mãos em repouso), a ocorrência da fase indica o interesse do então-receptor de dar início a um turno de fala; já quando a fase de preparação antecede um sinal *não-inicial* numa dada unidade gestual (i.e. a preparação de um sinal em meio a uma cadeia de sinais), a ocorrência da fase deve indicar o interesse do falante-corrente de dar continuidade à sua fala.

Nas LOs, a fase de preparação do gesto não tem praticamente recebido atenção dos analistas, em parte porque a preparação dos gestos articulatórios de uma palavra oral ocorre no interior do trato vocal, com pouca ou nenhuma manifestação corporal observável ao interlocutor; e em parte porque a atividade muscular envolvida na realização dos gestos articulatórios de palavras orais parece ser significativamente menor do que a atividade envolvida na preparação de sinais em línguas como a libras. Apesar disso, a discussão de Schegloff (2000) sobre a etapa de pré-inicialização das sobreposições revela que, mesmo nas LOs, tal fase também se mostra relevante para o fenômeno da sobreposição (p. 15).

Segundo o autor, antes de duas “falas” se encontrarem sobrepostas, é possível que os participantes se orientem a certas formas de gestualidade indicativas de que o interlocutor está para iniciar um turno. Schegloff cita como pistas dessa etapa “pré-inicial” das sobreposições o que ele chama de “sinais comportamentais do corpo” (i.e. gesticulações manuais, realinhamento postural) e “práticas comuns para iniciar turnos” (i.e. inspiração audível de ar). Em orientação a essas formas de gestualidade, recursos tais como o aumento de volume e a aceleração do ritmo da fala seriam empregados pelos falantes-correntes a fim de se prevenirem contra a ocorrência de uma “sobreposição de fato”.

Um exemplo extremo que o autor traz é o de uma conversa em que, diante de uma inspiração de ar do então-receptor, típica de inicializações de turnos, o falante-corrente responde: “Wait don’t interrupt me.” (p. 50). O exemplo é extremo porque, como o próprio autor argumenta, são raros os casos de sobreposição em que a resolução

é alcançada trazendo-se o próprio problema para o foco da interação, de maneira explícita, ao invés de abordá-lo implicitamente por meio de recursos e procedimentos convencionais aplicados à fala (i.e. aumento do volume, cortes abruptos, etc). A respeito da passagem, então, é interessante ver o modo como Schegloff opta por introduzi-la em sua discussão: “A instância dessa orientação [à gestualidade preparatória dos turnos] que aparece a seguir ... serve como um exemplo dramático, embora possa parecer inadequado aqui pelo fato de *nenhuma sobreposição de fato* ocorrer na passagem” (p. 50).

Tal afirmação mostra que, para Schegloff, os conceitos de *fala* e/ou *sobreposição de fala* estão restritos à dimensão verbal da língua. A relevância da fase de preparação do gesto, por exemplo, tende a ser ofuscada pela invisibilidade e/ou sutileza dos gestos envolvidos, ou ainda pela diferença entre a produção estritamente vocal e a produção manual/corporal que a acompanha. E mesmo a perspicácia do autor, capaz de identificar a orientação dos participantes em relação a essas fases preparatórias do gesto, não bastou para que o estatuto lingüístico dessas formas de gestualidade fosse colocado em questão. Nas LSs, diferentemente, a grande saliência da fase de preparação dos gestos articulatorios, aliada ao fato de que os gestos opcionais são produzidos pelo mesmo canal de “fala” sinalizado, fazem com que essa dimensão gestual salte aos olhos em análises da fala-em-interação.

A reflexão acima sobre as fases de preparação e retração dos gestos sugere, portanto, que análises sobre o contraste entre sobreposições de fala em línguas de diferentes modalidades tendam a estar duplamente enviesadas: nos estudos das LSs, fases como a retração parcial e total de sinais em sobreposição tendem a ser *incluídas* sob o escopo de “fala sobreposta”, o que, especialmente no caso da retração parcial, aumenta significativamente a extensão do fenômeno; e, nos estudos das LOs, fases como a preparação, seja de palavras/frases, seja de gestos concomitantes à fala na inicialização de turnos, tendem a ser *excluídas* do escopo de “fala sobreposta”, o que diminui a extensão do fenômeno.

Na passagem analisada nesta seção, por exemplo, vimos que a duração da sobreposição de fala sem uma consideração das fases do gesto é de 2s135ms. Se excluirmos as *retrações parciais e totais* ao final das unidades (327ms e 390ms respectivamente) por não se mostrarem seqüencialmente implicativas, como argumentado acima, a duração dessa instância cai para 1s418ms. E se excluirmos a fase

de preparação ao início do turno (302ms), como os estudos das LOs têm implicitamente feito, a duração da mesma ocorrência cai ainda mais, para 1s116ms.

Além disso, embora o aceno de cabeça de W tenha sido aqui considerado um item lexical pleno, representável na transcrição por meio da glosa SIM (figura 22, trilha W.3, ver discussão na nota de rodapé 148), há formas na libras de se responder afirmativamente a uma pergunta utilizando sinais manuais (ou combinações de sinais manuais e não-manuais). Até que ponto esse sinal não-manual, então, deveria ser contado como uma fala plena? Uma forte evidência de que deva é a posição na seqüência interacional em que esse aceno ocorre, aparecendo como resposta direta à questão de R e envolvendo, portanto, uma forte implicação seqüencial na situação imediata da interação. Contudo, nos estudos de LOs, é bastante provável que uma resposta afirmativa a uma questão por meio de um aceno de cabeça desacompanhado de vocalização não fosse contada como sobreposição – seja pelo fato de o analista não ter acesso à imagem da interação, seja pelo fato de o analista não considerar “gestos” com o mesmo estatuto da “fala”. Desse modo, se seguíssemos novamente a orientação geral dos estudos das LOs, o período relativo a esse aceno (712 ms) também deveria ser excluído do cômputo da sobreposição, reduzindo a ocorrência a meros 404ms – 1s731ms a menos do que apontava a análise inicial.¹⁴³

A análise acima mostra que a menos que critérios analíticos bastante claros para a segmentação da cadeia da fala sejam definidos para ambas as modalidades de língua,

¹⁴³ A questão sobre o que considerar fala plena e, conseqüentemente, sobreposição de fala problemática, traz ainda outras complexidades. Por exemplo, na análise acima, considerei, a título de reflexão, a exclusão da preparação tanto dos sinais não-manuais quanto do primeiro sinal na unidade gestual, PROFESSOR, mas não disse nada a respeito das preparações dos sinais subseqüentes nessa unidade (e.g. figura 22, trilha R.3, “pr” sob os sinais EXPLICAR e VOCÊ). Diferentemente do primeiro sinal ao início de um turno, cuja preparação é bem mais saliente pelo fato de as mãos terem que se deslocar do repouso até o espaço de sinalização, a preparação dos sinais subseqüentes numa unidade gestual integram uma cadeia de fala fluente, envolvendo apenas o investimento motor necessário para fazer a transição de um sinal para outro. Tendo em vista que essas transições fluidas entre palavras não são excluídas das análises de sobreposições em LOs, porém, não seria pertinente, neste exercício de reflexão, excluí-las aqui da análise da conversa em libras.

Essa observação nos remete à discussão anterior sobre a saliência da fase de preparação nas LSs, em comparação com a das LOs. É evidente que a preparação dos sinais deva envolver, seja no início do turno, seja em meio a uma cadeia de fala fluente, um investimento motor bem maior do que a preparação de palavras nas LOs. É plausível especular, portanto, que essa diferença acarrete, objetivamente, tempos maiores de sobreposição nas conversas sinalizadas quando comparadas às conversas em língua oral. Se isso acontecer, porém, a motivação desses períodos maiores de sobreposição não deve ser atribuída, como tem sido equivocadamente feito, à capacidade de produzir e compreender fala simultaneamente nas conversas em LSs, mas às diferenças motoras de produção em línguas de modalidade distintas.

quaisquer generalizações sobre diferenças e/ou semelhanças entre o sistema de troca de turnos em LOs e LSs devem ser tomadas como potencialmente equivocadas. Em particular, tendo em vista o fato de que trabalhos como os de Baker (1977) e Coates e Sutton-Spence (2001) não levam tais questões em consideração, suas conclusões (i.e. de que “surdos falem mais em sobreposição do que ouvintes”, de que “surdos possam produzir e compreender a fala simultaneamente”, ou de que “todos os participantes surdos atendam a todos os demais ao mesmo tempo”) *não* devem ser tomadas como uma evidência teórica e metodologicamente robusta para a rejeição do modelo do um-de-cada-vez.

6.4.2. Circunstâncias de fala simultânea excluídas

Como discutido na seção 6.2 acima, pelo menos três situações de sobreposição de falas são tidas como convencionalmente não-problemáticas para os participantes: as sobreposições terminais (i.e. quando um turno se inicia no momento em que o outro está bastante próximo de alcançar o fim); as sobreposições com turnos minimalistas (i.e. quando um turno em curso é acompanhado, por exemplo, por continuadores, cuja função é a de ratificar um turno mais extenso por parte do falante-corrente); e as sobreposições que fazem parte de práticas convencionais que, para serem realizadas de maneira adequada, envolvem diferentes tipos de participação simultânea (e.g. risadas).¹⁴⁴

A título de ilustração, passagens referentes às duas primeiras situações foram identificadas no corpus e analisadas. A primeira dessas passagens, que aparece na figura 23, refere-se a <[uma instância de sobreposição terminal](#)>. No excerto, W está reclamando do fato de, na avaliação de física do seu grupo, ele ter ficado responsável pela explicação de seis conceitos, ao passo que o seu colega (cujo sinal/nome foi glosado como I-L2, trilha W.2) ficou com apenas dois. Frente a essa reclamação, R inicia um turno para argumentar que nem todas as pessoas têm o mesmo gosto e que, no caso de alguém que não gosta de física, fazer algo simples já basta.

¹⁴⁴ Como já apontado, a outra forma de sobreposição não-problemática, quando dois participantes falam simultaneamente em conversas próximas porém separadas, não pôde ser observada nos dados pelo número restrito de participantes nas gravações.



Figura 23. Exemplo de sobreposição terminal entre turnos na conversa em libras.

A passagem mostra que o turno-em-curso de W segue com uma boa fluência até a frase “Ele só com dois!” (à esquerda da faixa vertical escura), onde aparece pela primeira vez uma retração parcial (trilha W.3). Nesse ponto, em ótima coordenação com W, cerca 0,1 segundo após o início da retração de DOIS, R inicia a preparação de um gesto muito comum na libras, em que a mão é levada à altura do peito com a palma voltada para cima. Esse gesto é acompanhado de um conjunto de sinas não-manuais: encolhimento dos ombros, inclinação da cabeça para o lado e, no rosto, levantamento da sobrancelha, fechamento parcial das pálpebras e abertura da boca – ver figura 24.¹⁴⁵ A meu ver, essas formas de gestualidade seriam o tipo de fenômeno ao qual Schegloff (2000) se refere pelo uso do termo “sinais comportamentais do corpo”, que envolvem aspectos tais como gesticulações manuais e realinhamento da postura e revelam o desejo do participante de tomar o turno.¹⁴⁶

¹⁴⁵ A mão direita de R realiza um gesto de arrumação do cabelo que não me parece relevante neste contexto e, portanto, deve ser desconsiderado.

¹⁴⁶ Minha impressão é a de que, embora tais gestos, em combinação com os não-manuais, possam ser uma “demonstração do desejo de tomar um turno”, seria equivocado restringir o seu sentido apenas a essa função. Na passagem em questão, por exemplo, optei por glosar o gesto manual por AH (trilha R.2), pelo fato de, intuitivamente, a preparação da participação de R se assemelhar à preparação que nós, ouvintes, fazemos no português, em frases como: “Ah, depende de cada pessoa!” – cujo significado eu não saberia agora precisar, mas que certamente uma investigação sobre as suas várias instâncias poderia revelar.

Nesse sentido, duas observações – que servem tanto aos estudos de LSs quanto aos de LOs – podem ser feitas. Em primeiro lugar, é possível que determinadas formas de “gestualidade” que a princípio seriam associadas à “preparação para um turno” se mostrem altamente convencionizadas, expressando idéias precisas com tal sistematicidade que excluí-las do repertório lexical da língua seria uma decisão arbitrária; nesse sentido, o próprio estatuto não-lingüístico desses gestos, como “*preparatórios* para um turno”, deveriam ser colocados em questão (ver discussão na seção 6.4.1 acima). Em segundo lugar, mesmo gestos não convencionizados, tal como posturas corporais com níveis de alinhamento gradientes em relação ao interlocutor (i.e. inclinação para frente vs. para trás; orientação para esquerda vs. direita; posição relaxada/curvada vs. tensa/ereta) podem iconicamente representar graus de envolvimento, participação, interesse, atenção que vão além de meramente refletir um desejo “neutro” de tomar o turno.



Figura 24. Gestos manuais e não-manuais na preparação para a tomada do turno.

Praticamente no mesmo instante em que R inicia a preparação desse gesto, W opta por dar prosseguimento à sua fala por meio do sinal FOLGAD@. A percepção de que R resolve iniciar um turno no mesmo momento em que ele resolve continuar o seu, porém, resulta no abandono imediato da sua fala assim que ela alcança uma possível completude (por meio da retração total após o golpe do sinal FOLGAD@). Essa retração ocorre no exato momento em que R alcança o pico do gesto preparatório, AH, dando continuidade à sua fala (com a frase “Depende de cada pessoa”, trilha R.1, 37s100ms). Sobreposições terminais caracterizam-se precisamente por essas marcas: o abandono do turno-em-curso pelo falante-corrente tão logo a unidade em curso alcance a sua completude, com leve sobreposição de fala durante a transição entre falantes, e a emergência de um novo turno produzido em claro por um dos interlocutores da conversa. Nenhuma disfluência ou perturbação se mostra notável em circunstâncias como essa.¹⁴⁷

Uma segunda situação de sobreposição que convencionalmente não afeta a fala dos participantes é a simultaneidade entre um turno em curso e contribuições minimalistas do então-receptor, que não almeja o direito pela palavra – ver discussão nas seções 3.4 e 3.5. Um exemplo que aparece no corpus da libras é o do [<uso de continuadores por meio de leves e contínuos acenos de cabeça>](#). No excerto em questão, que sucede aquele discutido na figura 20, W está explicando à R a sua dificuldade e ansiedade com uma aula de física em que ele fora chamado para dar uma explicação ao professor. Em meio à narrativa de W, R intervém com um questionamento (“Você ia explicar para o professor?”), um reparo aparentemente para garantir o entendimento sobre a ansiedade que W demonstrava. Em sua resposta, W diz que seria avaliado, não apenas ele como todo o seu grupo, formado por quatro alunos, e, em seguida, retoma a

¹⁴⁷ Como será discutido na seção 6.4.3 abaixo, tanto a retração parcial quanto a total aparecem nos dados da libras como recursos para gerenciamento de sobreposições problemáticas. Quando isso acontece, porém, elas emergem em determinados pontos da fala em curso que apresentam uma clara incompletude, diferente do caso aqui analisado.

narrativa no exato ponto em que ele havia sido “interrompido” por R com esse questionamento – retomando inclusive as mesmas palavras que ele havia usado anteriormente. Durante cerca de nove segundos, em que W reporta essas outras informações em resposta ao reparo, R realiza acenos de cabeça bastante sutis e contínuos, que entendo serem instâncias de continuadores na conversa em libras.¹⁴⁸ Como esperado, esse nível de participação de R não traz quaisquer reflexos na fala de W, seja na forma de perturbações, seja de disfluências.

Junto com uma definição clara sobre a segmentação da cadeia de fala, a consideração sobre esses diferentes tipos de sobreposição não-problemáticos constitui um passo fundamental para qualquer comparação entre a natureza das sobreposições em LSs e em LOs. Desconsiderá-las significa incluir como parte do fenômeno uma série de ocorrências que, a princípio, não são tratadas como problemáticas pelos participantes mesmo dentro da proposta de SSJ. Tendo em vista que nem Baker (1977) nem Coates e Sutton-Spence (2001) relacionam claramente as suas análises às diferentes circunstâncias de sobreposição, mais uma vez as suas generalizações a respeito de uma possível alternativa ao modelo do um-de-cada-vez se mostram enfraquecidas.

6.4.3. Recursos gestuais para o gerenciamento de sobreposições

Feitas as devidas considerações metodológicas e teóricas relativas ao processo de gerenciamento da troca de turnos e das sobreposições na libras, podemos então passar às circunstâncias de sobreposição – predominantes no corpus – que se mostram problemáticas aos participantes. Na medida em que a fala dos participantes nessas passagens apresente disfluências e perturbações (i.e. recursos de resolução de sobreposições), tal análise deverá se constituir numa fonte de evidência de que os participantes surdos na conversa em libras também se orientam em relação ao modelo do um-de-cada-vez.

¹⁴⁸ Uma observação cuidadosa da dinâmica desses continuadores, em contraste com o aceno de cabeça de W em resposta à R na passagem da figura 20, sugere que uma diferenciação entre tipos de aceno com funções interacionais distintas possa ser estipulada. Tal diferenciação, contudo, não deve ser tomada como um argumento contra a natureza gradiente desses gestos. Analisando esses acenos em termos de sua iconicidade, por exemplo, podemos observar que quanto maior o deslocamento e a duração das fases gestuais do aceno, maior o nível de participação/envolvimento na interação e, conseqüentemente, maior a implicação seqüencial do aceno – da implicação de um item lexical até a de um continuador (ver discussão sobre níveis de envolvimento na interação, na seção 3.5).

De fato, a análise dos momentos em que mais do que um fala ao mesmo tempo na conversa em libras revela o uso sistemático de certos recursos visando à resolução das sobreposições de fala. Na maioria dos casos, tais recursos aparecem em combinação uns com os outros numa única ocorrência. Os recursos identificados nos dados são:

- Reciclagens da fala sobreposta, que podem envolver um único sinal, uma expressão dentro de uma mesma unidade, unidades inteiras, ou expressões e unidades com pequenas modificações;
- Interrupção abrupta da fala por meio de realização incompleta do sinal e/ou unidade, seguida de retração total ou de retração parcial, com fragmentos do sinal podendo pairar no ar por um período prolongado de tempo;
- Aumento do volume do sinal, por meio de uma realização mais ampla, isto é, um deslocamento maior no espaço;
- Desaceleração da fala por meio de alongamentos prolongados, com os sinais apresentando manutenção da suspensão independente ou reiteração do golpe dependendo do tipo de sinal alongado;
- Aceleração da fala para a unidade seguinte, revelada por meio de uma antecipação marcada dos sinais não-manuais referentes ao sinal e/ou à unidade subsequente;
- Fechamento prolongado dos olhos e olhar divergente;
- Pedidos explícitos de atenção por meio de sinais reguladores (e.g. acenos e itens lexicais como FICAR), às vezes interpostos aos sinais nas frases, às vezes concomitantes à sinalização;
- Itens lexicais para o licenciamento de turnos extensos, tais como OLHAR e EXEMPLO.

O excerto que aparece na figura 25 oferece uma primeira ilustração da combinação de alguns desses recursos. <[Nessa passagem](#)>, W está começando a reportar à R sobre uma aula em que ele havia sido chamado para dar explicações para a turma, momento em que R inicia um reparo questionando-o sobre o conteúdo dessa explicação (EXPLICAR O-QUE, trilha R.2). W parece antecipar o questionamento de R, pois imediatamente ao término da preparação do sinal EXPLICAR (“pr/st”, trilha R.3, 5s), W já inicia a preparação de sua resposta, FÍSICA (trilha W.2). Essa resposta, que emerge em sobreposição com a fala de R, é reciclada três vezes antes de R

demonstrar reconhecimento da resposta por meio do sinal FÍSICA (trilha R.2, 6s500ms).



Figura 25. Reciclagens sucessivas do sinal FÍSICA numa instância de sobreposição de fala.

A ocorrência de *reciclagens* do mesmo sinal numa única instância de sobreposição oferece a possibilidade de uma comparação formal entre sinais idênticos num mesmo contexto. Observando a dinâmica desses sinais, então, nota-se uma modulação relativa à extensão do deslocamento do sinal no espaço. Como mostra a figura 26, na primeira ocorrência de FÍSICA em resposta à questão de R, o sinal é realizado de maneira bastante breve, dando a impressão inclusive de ter a sua trajetória completa interrompida; na segunda, há um deslocamento significativamente maior; e na terceira, o deslocamento é intermediário entre o da primeira e da segunda ocorrência.

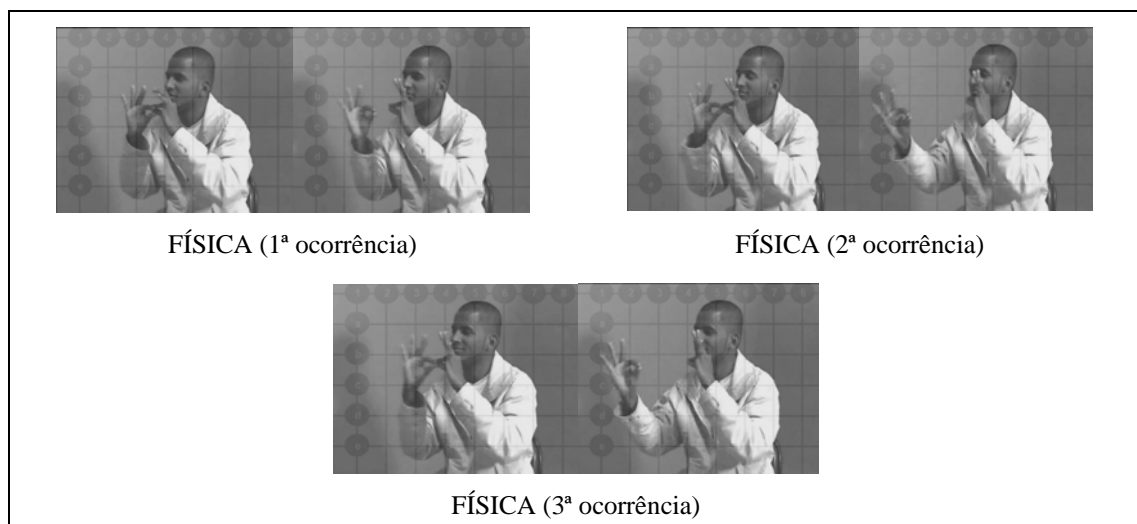


Figura 26. Variações de volume do sinal FÍSICA reciclado em sobreposição.

Nesse exemplo, pode-se especular que o aumento do volume do sinal tenha sido motivado pela necessidade de produzir uma elocução mais saliente da palavra, diante da permanência da dúvida de R ao longo das sucessivas reciclagens do sinal FÍSICA. Em outras ocorrências, porém, como a que aparece na figura 27, o aumento do volume

parece servir como um claro recurso de reivindicação da palavra. <Nesse excerto>, W estava dando início a um novo turno quando R recicla o sinal I-L2 (trilha R.2, após os 2m13s) com um claro aumento de volume.¹⁴⁹ Nesse momento, W *interrompe bruscamente o seu turno* em meio à preparação de um novo sinal não-identificado (P=, trilha W.2), que é seguido então de uma longa *retração parcial* intercalada por um gesto sutil, AH (trilha W.2, faixa vertical escura), e uma expressão facial demonstrando empatia frente à fala de R.¹⁵⁰

	00:02:11.000	00:02:12.000	00:02:13.000	00:02:14.000	00:02:15.000	00:02:16.000	00:02:17.000	00:02:18.000	00:02:19.000	00:02:20.000			
W.1		(Não) tenho.	Eu sentava p=		Ah...		Sente...(que) não (faz)	Isso	(Foi) expondo (o q) Eu.				
W.2		NÃO-TER	SENTA P=		AH		SENTIR	COM N	ISS	IS PALESTRAR EU			
W.3		pr st	pr st h pr	rt		pr rt		pr st	pr st pr	pr pr st pr rt			
W.4		ad	ri ad		ad		ad	ri	ad hd	ri le			
R.1	EH... você tem intérprete?	Nada a	(O) I-L2=	(O) I-L2 (não tem) culpa		Ele sente hm (que não faz)	C= (o) professor (é) complicado.		Ele não (tem) culpa.				
R.2	INTÉRPRETE	TER	NA AH I-L2=	I-L2	CULPA	EL AH	SENTIR	MM COMBMM	PES	PRO PESADO MM	AH EL@ CULP NÃO		
R.3	pr st	pr st	pr s pr	pr st pr	pr st	pr pr	pr st	pr pr st pr	pr p st pr st	pr p st pr st pr pr pr pr s pr rt			
R.4	ad		ad	ad		ad	le	ad	le	le	ad		ad

Figura 27. Recursos de reivindicação e de cessão da palavra combinados numa instância de sobreposição.

A coordenação entre a disputa da palavra por R, através da exploração do aumento do volume do sinal, e a cessão da palavra por W, por meio da interrupção da sinalização seguida de retração parcial, aparece nas imagens da figura 28. Ela mostra que, durante a realização do primeiro sinal I-L2, W ainda fala em sobreposição com R (EU SENTAR, trilha W.2, 2m12s700ms); entretanto, diante da reciclagem de I-L2 com o volume aumentado, W interrompe a sua sinalização, realizando uma retração parcial e juntando as mãos na altura do abdômen.

¹⁴⁹ As trilhas R.4 e W.4, referentes à direção do olhar, envolvem as seguintes siglas de notação: “ad” (*addressee*), “hd” (*hands*), “le” (*left*), “dn” (*down*).

¹⁵⁰ Vale lembrar que, na transcrição, temos optado por anotar toda a cadeia gestual, desde o início da preparação até o final da retração, o que gera a impressão de sobreposições maiores do que de fato são. Nesse excerto, por exemplo, o período referente a toda a faixa vertical escura, em que a fala de W envolve uma retração parcial e um gesto discreto, revelam na verdade o fim de uma sobreposição de fala bastante breve: quando W sinaliza “EU SENTAR P=” e R sinaliza “I-L2 IL-2”.

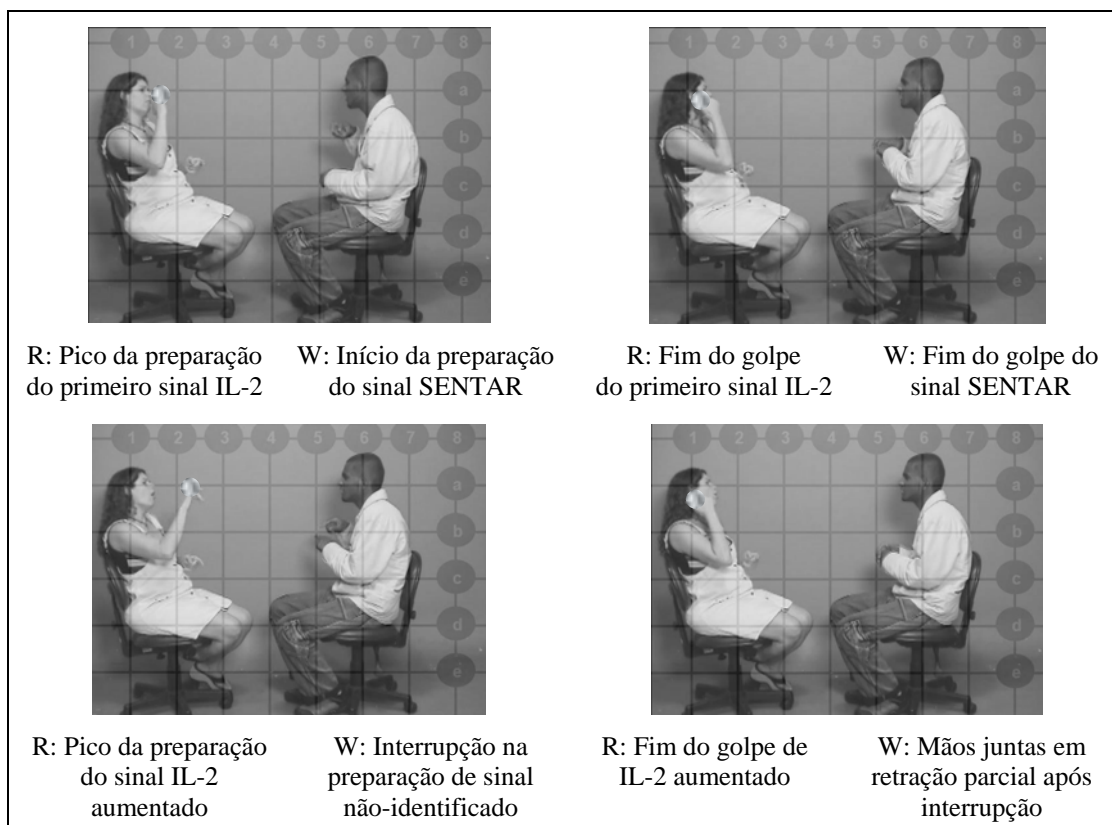


Figura 28. Eficácia do aumento do volume e da retração parcial na resolução de sobreposição de fala.

Ao término dessa resolução, W volta a falar em sobreposição com R quando repete a frase de sua interlocutora, “Sente (que) não (faz) sentido” (W.1, 2m17s e R.1, 2m15s200ms). Analisando essa retomada da fala por W, nota-se que ele emprega aqui dois recursos para sinalizar o seu interesse pela palavra: primeiramente, o *alongamento de um sinal* e, em seguida, o *rompimento do contato visual* com R. Como mostra a figura 27, quando W inicia o sinal SENTIR (trilha W.2, ao término da faixa escura), o olhar de R está voltado para a esquerda (“le”, trilha R.4). Nesse ponto, então, W altera a progressividade de sua fala, alongando o sinal SENTIR por meio da reiteração do golpe, somente passando à realização do sinal subsequente, COMBINAR, quando consegue o re-estabelecimento do contato visual de R (“ad”, trilha R.4, 2m17s700ms). Tendo sido bem sucedido nessa primeira etapa, então, W explora em seguida a divergência do olhar (“hd”, “ri”, “le”, trilha W.4, 2m18s700ms) a fim de levar R ao abandono de seu turno, agora em sobreposição. Com a mesma eficácia que o alongamento de SENTIR, o olhar divergente de W repercute diretamente na fala de R, que parte para uma *retração total* em meio à realização do sinal NÃO, indicando um abandono completo do turno (“rt”, trilha R.3, 2m20s).

Tendo em vista o papel fundamental do olhar na sinalização do direcionamento e do nível atencional dos participantes – importância que parece ser ainda maior no caso de conversas em LSs em comparação com conversas em língua oral –, o olhar divergente e o fechamento dos olhos aparecem como recursos produtivos na disputa pela palavra. No <[excerto a seguir](#)>, na figura 29, R está terminando um turno por meio do qual avalia a aparente falta de interesse do colega de W pela matéria de física.¹⁵¹ Nesse ponto, em sobreposição com a última unidade de R, W inicia seu turno *fechando os olhos* por um período prolongado de tempo (“cb”, trilha W.4, durando doze frames¹⁵² ou 0,4 segundos). O fato de esse fechamento ser acompanhado por uma *longa retenção da suspensão independente* do sinal EU (“hi”, trilha W.3, durando sete frames ou 0,2 segundos) sugere que a estratégia seja profícua em circunstâncias de disputa pela palavra nas quais o participante deseja falar – porque o momento na seqüência interacional assim exige – mas ainda não tem totalmente planejada a fala a ser proferida.

	00:00:46.000	00:00:47.000	00:00:48.000	00:00:49.000	00:00:50.000
W.1	Eu... deixo.		Eu respeito (el	Você (é) folgado	Só dois! (Muito) pouco!
W.2	EU	PERMITIR	EU	RESPEIT	VOC FOLGAD DOIS
W.3	pr	hi	pr	h st	pr
W.4	cb		sb	sb	sb
R.1	simples.	Dois conceitos.	Depende (de cada) pessoa.		
R.2	MPLES	PALAV	PALAVRA?	DEPENDER	PESSOA

Figura 29. Fechamento prolongado dos olhos e retenção da suspensão independente de sinais sem golpe para garantir a palavra.

O <[excerto abaixo](#)>, na figura 30 demonstra uma instância na qual um dos participantes, W, utiliza o fechamento prolongado dos olhos em combinação com vários outros recursos para ganhar direito ao turno. A transcrição mostra que W inicia sua

¹⁵¹ A trilha W.4, referente à palpebra, envolve as seguintes siglas de notação: “cb” (*both eyelids closed*); “sb” (*narrowed eyelids*).

¹⁵² Cada frame num arquivo digital de vídeo corresponde a 33 milissegundos – ou, posto de outra maneira, 3 frames correspondem a 0,1 segundo e 30 frames a 1 segundo. Pelas observações feitas até agora, a medida de 3 frames parece ser uma boa referência para identificação de alongamentos finais.

Embora 0,1 segundo pareça ser um valor irrisório, tomei a decisão de considerar tal medida um alongamento com base na observação de que as *suspensões finais* dos sinais em posição inicial ou medial da frase ou são elididas, ou reduzem-se a não mais do que um único frame. A exceção para esse padrão observado parece ser os sinais que aparecem em soletrações manuais, que, como será visto a seguir, apresentam o padrão de três frames para cada sinal soletrado, o alongamento final se estendendo por um período maior, de cinco a seis frames.

contribuição por meio de um aceno com a mão dominante, momento em que a pálpebra está em posição neutra (trilha W.3, entre 1m a 1m00s600ms, sem anotação). Esse aceno é seguido da frase EU ACHAR (trilha W.2), acompanhada de uma abertura mais ampla dos olhos (“wb”, trilha W.3). O uso dos recursos até então parece não surtir o efeito desejado, uma vez que R retém o sinal VOCÊ no ar com uma longa suspensão independente (“hi”, trilha R.3), mantendo também a expressão não-manual de perplexidade que acompanha a sua pergunta, “Você... vai obrigar, você?!”. Diante de uma possível disputa pelo turno, W muda, então, a sua estratégia: ele fecha os olhos por um período de tempo prolongado (“cb”, trilha W.3, por 22 frames ou 0,7 segundos) e utiliza *sinais de licenciamento de turnos extensos* (OLHAR, EXEMPLO).¹⁵³ O final da reivindicação de W só ocorre quando, após *reciclar sua frase inicial com pequenas modificações*, “Eu acho” (trilha W.1, 1m2s100ms), W observa uma retração total no gesto final de R (“rt”, trilha R.3, 1m2s900ms).

	00:01:00.000	00:01:00.500	00:01:01.000	00:01:01.500	00:01:02.000	00:01:02.500	00:01:03.000	00:01:03.500	00:01:04.000	00:01:04.500			
W.1	Hey, esperal	Espera eu acho.	(Por) exemp= ó	(Por) exemplo.	Eu acho			Ela sentada,	A mulher ouvinte,				
W.2	HEI-2	EU	ACHAR	EXE	OLHAR	EXEMPLO	EU	PENSAR	ACHAR	EL@	SENTAR	MULHER	OUVI
W.3		wb	c	cb	sb	cb	sb	sb	wb				sb
R.1	Você... (vai) obrigar, você?!					Ah...							
R.2	VO	OBRIGAR	VOCÊ			AH							
R.3	pr	pr	st	ha	pr	hi	rt	pr	hi	rt			

Figura 30. Combinação de vários recursos para resolução de sobreposição na disputa pela palavra.

Um recurso de resolução de sobreposições ainda não discutido que W utiliza nessa passagem é a *exploração da mão não-dominante, simultaneamente à sinalização da mão dominante, para pedido explícito de atenção*. Como mostram as imagens da figura 31, enquanto a mão dominante de W (a direita) realiza um aceno – glosado como HEI-2 por ser formalmente distinto do aceno HEI – seguido da frase EU ACHAR, a mão não-dominante produz o sinal FICAR, ora realizado de maneira plena (i.e. com

¹⁵³ Ainda no excerto relativo à figura 30, é interessante notar que, num primeiro instante de breve abertura dos olhos (“sb”, trilha W.3, 1m1s500ms), W se depara com R sem uma alteração comportamental significativa em relação ao início da disputa, o que o leva a fechar novamente os olhos; somente no segundo momento de abertura, então, é que, observando R com um sinal que parece ser típico de inícios e finais de turno (AH, trilha R.2, 1m2s900ms), W volta a abrir os olhos de maneira definitiva (“sb” e “wb”, trilha W.3, 1m2s010ms).

golpe), ora simplificada (i.e. envolvendo apenas a preparação do sinal e a manutenção da sua suspensão final).¹⁵⁴



Figura 31. O uso simultâneo da mão não-dominante por meio do sinal FICAR na resolução de falas em sobreposição.

A análise do <[excerto a seguir](#)>, na figura 32, mostra que o fato de o interlocutor apresentar um olhar convergente ou divergente é determinante na escolha desses diferentes tipos de sinais reguladores. Na passagem que antecede a sobreposição, W está argumentando que o fato de seu colega ter menos conceitos do que ele para apresentar seria ruim, pois a professora poderia passar a exigir pouco de seu colega. É nesse momento que R pede a atenção de W para contradizê-lo, buscando licenciar um turno mais extenso por meio do sinal EXEMPLO, reciclado uma vez (trilha R.2, 1m11s).

	00:01:06.000	00:01:07.000	00:01:08.000	00:01:09.000	00:01:10.000	00:01:11.000	00:01:12.000	
W.1	Parece (que o) Wilson estud		Desenvolve (a) leitura,		Ó (que) bom!		Já (o) H2 parece (ruim).	
W.2	ACHAR	WML	ESTUDAR	LER	DESENVOLV	VER	BOM	
W.3	pr	s	st	p	st	pr	st	
W.4	ad	dn	dn	hd	le	le		
R.1					Hei (por) exemplo.	Hei (por) exe	Você.	Se (no) futur
R.2					HEI	EXEMPL	FIC	EXEMPL
R.3					pr	st	pr	st
R.4					ad	ad	ad	ad

Figura 32. A dinâmica do olhar na resolução de uma sobreposição de fala.

¹⁵⁴ À primeira vista, o sinal em questão poderia ser interpretado como um gesto com a função específica de regular a troca de turnos. Contudo, a forma similar à do item lexical pleno, FICAR – em contextos da libras como <“[Imagine... se você ficasse lá \[na sua cidade natal\], faria a mesma coisa](#)”> – sugere que se trate do mesmo sinal, que, em situações de sobreposição de fala, adquire um sentido similar ao do português “esperar”, em frases como: “*Espera*, eu acho que...”. Nesse sentido, o item lexical FICAR, assim como os itens OLHAR e EXEMPLO no licenciamento de turnos extensos, revelam uma função reguladora na troca de turnos que emerge como extensão de sua significação lexical. Essa observação vem ilustrar o argumento, discutido nas seções 4.2 e 4.3, sobre a impossibilidade de se cindir categoricamente a função referencial e interacional da língua, como Chafe parece fazer.

A pequena sobreposição de fala envolve pedidos explícitos de atenção por meio do uso de dois sinais reguladores (HEI e FICAR). Quando R produz o primeiro sinal regulador, a fala de W está assumindo a voz e perspectiva da professora, razão pela qual seu olhar, divergente, se mostra direcionado à esquerda (“le”, trilha W.4, 1m9s750ms); ao início do segundo sinal, porém, R já conseguiu o contato visual de W (“ad”, trilha W.4). Essa diferença entre a divergência e a convergência do olhar se reflete, respectivamente, no uso de HEI *vs.* FICAR, que se distinguem formalmente pelo tipo de movimento (i.e. um movimento local das mãos para cima e para baixo em relação ao punho no caso de HEI, e um golpe ou suspensão da mão espalmada fixa à frente, no caso de FICAR) e pela amplitude do deslocamento do braço no espaço de sinalização, como mostra a figura 33.

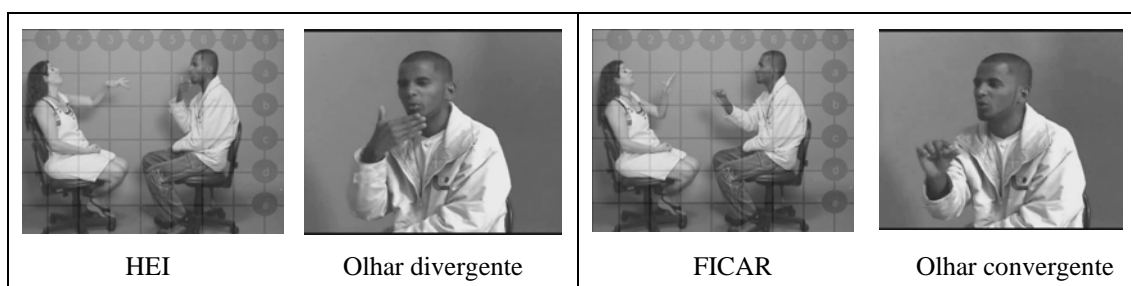


Figura 33. Correlação entre divergência/convergência do olhar e a escolha de sinais reguladores em situações de sobreposições de fala.

Duas observações referentes ao papel do olhar e ao papel dos sinais reguladores podem ser feitas sobre essa passagem. Sobre o olhar, nota-se claramente, como afirma Baker (1977) e McIlvenny (1995), a importância fundamental do contato visual para o estabelecimento da transição entre falantes na conversa sinalizada. Por exemplo, contrastando esse excerto com a ocorrência de “sobreposição terminal” – figura 23, seção 6.4.2 – nota-se que, em ambos, a sobreposição ocorre pelo fato de os dois interlocutores iniciarem uma nova contribuição ao término de uma unidade prévia de W. O estatuto problemático da atual passagem já ao início do turno de R, então, em contraposição ao estatuto não-problemático do excerto da figura 23, parece estar relacionado pelo menos em parte ao papel do olhar. No caso da sobreposição terminal não-problemática, R tem o contato visual de W, de modo que a preparação do seu turno (por meio do sinal AH) é diretamente monitorada por W, o que possibilita uma coordenação fina entre o término da preparação do turno de R e o início da retração da fala de W; já no atual excerto, o olhar divergente de W impossibilita essa coordenação

final, o que torna necessário o aceno inicial de R para a captação do olhar de W, e, em seguida, a reciclagem do sinal EXEMPLO para a garantia do licenciamento de um turno extenso.

Ao mesmo tempo, o fato de o olhar convergente ser uma condição *necessária* para a efetivação da transição entre falantes não implica que ele constitua condição *suficiente* para isso. Embora Baker (1977: 221) esteja correta ao afirmar que vários pedidos explícitos de atenção sejam resultado direto da necessidade de contato visual para a troca de turnos na conversa em ASL, a presente análise mostra que nem todos os sinais desse tipo podem ser explicados dessa maneira. No excerto em questão, por exemplo, R emprega o sinal regulador FICAR mesmo após a obtenção do olhar de W (“ad”, trilha W.4, aos 1m10s500ms). O motivo é o de que a dimensão perceptual da troca de turnos – para onde olha o interlocutor a cada novo momento – é apenas um dos fatores relevantes para indicar *o direcionamento e o nível atencional* dos participantes, esse sim o elemento crucial para a organização do um-de-cada-vez. Outro fator relevante é o fato (aparentemente trivial) de W continuar sinalizando, o que revela a sua orientação ao planejamento e execução de sua própria fala, em detrimento de uma orientação à contribuição que R deseja fazer. A retomada do contato visual, portanto, não basta para que R tenha assegurada a atenção plena de W; para garanti-la, R se vê na necessidade de empregar um segundo pedido explícito de atenção, FICAR, e de reciclar o sinal EXEMPLO. Somente então, diante da retração total na fala de W (trilha W.3, aos 1m10s900ms), é que a fala de R deixa de ser desenhada especificamente para a reivindicação da palavra e passa a tratar da contribuição de seu turno propriamente dita.¹⁵⁵

Outras instâncias do corpus em que esses dois sinais reguladores aparecem corroboram a hipótese de que eles estejam diferentemente correlacionados com a

¹⁵⁵ É interessante pensar que, na conversa sinalizada, os participantes rotineiramente se utilizem de falas explícita e separadamente devotadas à resolução de situações de sobreposição, tendo em vista a afirmação de Schegloff (2000: 51) de que, nas LOs, o problema da sobreposição é quase sempre tratado de maneira implícita, por meio de modificações na produção das próprias contribuições que emergem em sobreposição. Uma razão dessa diferença, como Baker aponta, certamente pode ser atribuída aos pedidos explícitos de atenção necessários ao estabelecimento do contato visual quando o interlocutor apresenta um olhar divergente; contudo, como visto, outros tipos de pedidos dessa natureza são freqüentemente empregados mesmo diante do estabelecimento do contato visual.

Uma primeira investigação, necessária para evitar a postulação prematura de diferenças entre conversas em LOs e LSs, seria a de verificar em que medida as sobreposições de fala nas conversas em LOs apresentariam formas similares de uso da gestualidade.

ausência ou presença do contato visual. Por exemplo, no <excerto a seguir>, na figura 34, os pedidos explícitos de atenção envolvem os mesmos dois sinais acima discutidos, HEI e FICAR. Quando R inicia seu turno, W está reportando diretamente um diálogo hipotético dele próprio com seu colega, razão pela qual seu olhar está direcionado à esquerda (“le”, trilha W.3, 50s650). R tenta iniciar o turno sem qualquer pedido explícito de atenção, mas diante da manutenção do olhar divergente de W, opta por fazer um aceno (HEI) e reciclar o sinal VOCÊ na frase: “Você (vai) obrigar (os) dois (a serem) iguais?” (trilha R.2, 51s800ms). Após uma breve troca em que W insiste em exigir que seu colega seja igual a ele, repetindo a frase “Eu quero igual!”, R aproveita o contato visual de W (“ad”, trilha W.3), para fazer uma nova solicitação da palavra, realizando o sinal FICAR acompanhado do sinal EXEMPLO, que vem mais uma vez licenciar um turno extenso, acarretando uma retração total na fala de W.

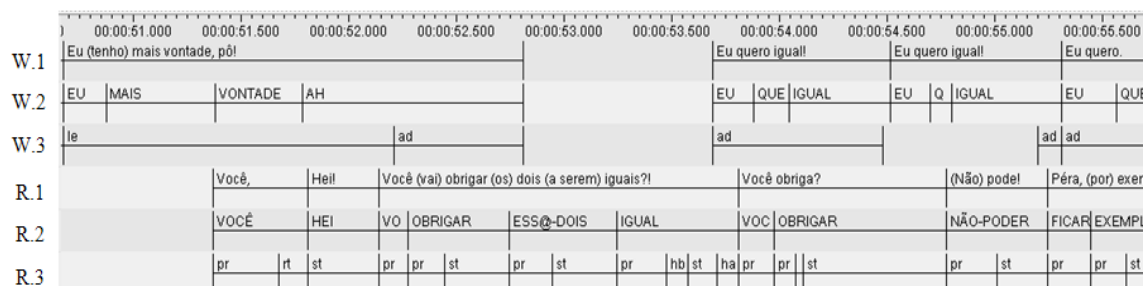


Figura 34. Pedidos explícitos de atenção diante da convergência e da divergência do olhar.

A passagem revela ainda o que eu interpreto ser uma *aceleração na sinalização* de R a partir do momento em que, diante das respostas sucessivas de W contrariando diretamente o seu argumento, R dá continuidade à sua fala buscando reforçar o seu ponto. Curiosamente, essa aceleração não se revela pela dinâmica dos próprios sinais manuais, mas pelo rompimento da coordenação desses com os sinais não-manuais ao final das frases “VOCÊ OBRIGAR” e “NÃO-PODER”. De maneira geral, os dados do corpus mostram que há uma coordenação fina entre os sinais manuais e os não-manuais característicos de certos tipos de orações (e.g. interrogativas e negativas), de modo que é comum uma precisão exata no início das preparações manual e não-manual, em alguns poucos casos a preparação não-manual se antecipando à manual em no máximo três frames ou 0,1 segundo.¹⁵⁶ Nas duas frases em questão, porém, a preparação dos sinais

¹⁵⁶ Posner (1980: 346) aponta que os sistemas fisiológicos para o controle das mãos e do olhar, por exemplo, são distintos, e que os olhos frequentemente se movem em direção a um estímulo antecipando os movimentos manuais. É possível que haja razões fisiológicas, portanto, para a propensão do olhar,

não-manuais se destaca por se antecipar à dos sinais manuais em cinco frames, ou 0,25 segundos. A impressão ao observar essa forte antecipação é a de que R está se apressando para chegar à unidade seguinte (i.e. acelerando o ritmo de sua fala) de modo a garantir o seu direito à palavra.¹⁵⁷

A figura 35 traz as configurações dos sinais não-manuais que marcam as frases “Você obriga?” e “(Não) pode!” (imagens “a” e “c”), ao lado da expressão já modificada de R no momento de realização do golpe final dos sinais terminais em cada uma dessas unidades: OBRIGAR e NÃO-PODER (imagens “b” e “d”). No caso da interrogativa, “Você obriga?”, destaca-se o contraste entre a inclinação da cabeça à frente típica de perguntas sim/não (imagem “a”) e o movimento lateral da cabeça referente à frase negativa subsequente (imagem “b”). No caso da negativa, “Não pode!”, destaca-se o contraste entre o franzimento da sobrancelha e a protuberância dos lábios associadas à negação enfática (imagem “c”) e o levantamento da sobrancelha e inclinação da cabeça para trás que acompanham o pedido explícito de atenção da frase subsequente (imagem “d”).

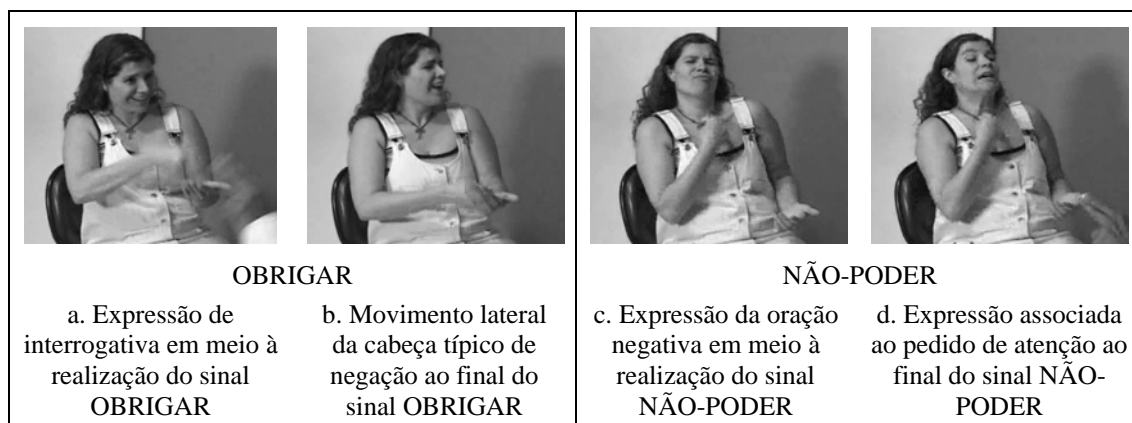


Figura 35. Aceleração da fala por meio de forte antecipação de sinais não-manuais relativos a sinais e/ou unidades subsequentes.

Como já afirmado, os recursos identificados nesta seção aparecem, na maioria das instâncias de sobreposição, combinados uns com os outros. Por esse motivo, em

juntamente com outros articuladores do rosto, de antecipar os sinais manuais com relativa frequência, em especial em situações nas quais o falante se apressa em produzir a unidade seguinte.

¹⁵⁷ Na literatura da AC, Schegloff (1982) identifica um fenômeno, que ele chama de *rushing-through*, segundo o qual um falante acelera a fala ao se aproximar do fim de uma unidade a fim de passar mais rapidamente o espaço de transição para a unidade seguinte. O autor argumenta que esse recurso tem como função sinalizar ao interlocutor que o falante-corrente pretende dar continuidade à sua fala, evitando a emergência de um turno em sobreposição.

alguns casos pode não ficar claro em que medida o sucesso na resolução da sobreposição se deve a um ou outro dentre os fatores presentes, ou se, na verdade, é a combinação sistemática e convencionalizada desses recursos que resulta numa resolução rápida e eficiente. Apesar da limitação do corpus utilizado como base para a presente investigação, a análise acima traz evidências suficientes para argumentarmos que a conversação na libras – e possivelmente em outras LSs – também trate como problemáticas as situações conversacionais em que se observa mais-do-que-um-de-cada-vez.

6.5. Conclusão

O presente capítulo envolveu uma análise sobre a troca de turnos na libras, com foco particular sobre as situações de transição entre falantes que resultam em sobreposição de fala. A escolha desse foco não foi arbitrária. Sem um conhecimento sólido sobre as unidades que constituem a fala em libras, simplesmente investigar as transições fluidas entre falantes (i.e. aquelas que não envolvem sobreposições ou intervalos) seria um procedimento analítico fortemente propenso à circularidade: a transição foi fluida porque se chegou ao fim de uma unidade; e chegou-se ao fim de uma unidade porque a transição foi fluida.

A estratégia alternativa para corroboração do modelo do um-de-cada-vez, então, foi a de identificar na conversa momentos em que tal princípio fosse claramente violado. Seguindo a proposta metodológica de Garfinkel (1967), assumiu-se que, se a conversação de fato opera sob esse princípio, a fala dos participantes não deveria passar incólume por situações em que ele fosse violado, de modo que disfluências e perturbações identificadas poderiam servir de evidência para o funcionamento do modelo.

Para a realização dessa análise, então, um primeiro passo foi o de olhar para como a questão da sobreposição de fala é resolvida nas LOs. Schegloff (2000) trata dos principais recursos e procedimentos envolvidos no trabalho interacional de resolução de situações problemáticas de sobreposição, constituindo-se, portanto, no meu ponto de partida da análise. A discussão do autor reforçou ainda alguns cuidados teórico-metodológicos (i.e. a exclusão de certas circunstâncias interacionais do escopo do fenômeno da sobreposição) que se mostram necessários para que o modelo do um-de-cada-vez possa ser apropriadamente avaliado.

Antes de me debruçar sobre o corpus da libras, os principais estudos que tomaram por objeto a conversação em LSs foram revisados com o intuito de apresentar as suas contribuições. Baker (1977) se aproxima do presente estudo pela sua preocupação empírica em descrever aquilo que ela denomina como “reguladores” da troca de turnos da conversa em ASL, o que serviu de importante base para a atual análise. Contudo, a afirmação da autora de que a conversa sinalizada permitiria a produção e observação da fala simultaneamente e sem perda de compreensão foi questionada, tanto com base nos próprios dados da autora (que mostram vários tipos de disfluências, perturbações e pedidos de atenção), quanto do argumento teórico desenvolvido no capítulo 4: a idéia de que a coordenação da ação social numa interação exija a atenção conjunta dos participantes sobre um mesmo foco a cada novo momento, algo que, tipicamente, é impossibilitado em situações de fala simultânea.

O trabalho de Coates e Sutton-Spence (2001) foi o que se revelou mais distante do presente estudo. Em minha leitura desse trabalho, procurei apontar alguns problemas teórico-metodológicos (i.e. inobservância de cuidados no arranjo das gravações e na escolha de excertos para análise, além de uma compreensão equivocada da proposta de SSJ) e empíricos (i.e. interpretação enviesada dos próprios dados apresentados). Tais problemas comprometem diretamente a afirmação das autoras de que a conversação na língua de sinais britânica, em especial entre mulheres, permita uma organização do tipo todos-atendem-a-todos-a-cada-momento.

McIlvenny (1995), diferentemente, se aproxima do presente estudo por sua orientação teórica, que, pautada na AC, pressupõe a pertinência do modelo do um-de-cada-vez na conversa sinalizada – embora o autor não faça uma análise para comprová-la. Seu foco analítico foi sobre a identificação de práticas no gerenciamento da troca de turnos na interação sinalizada (i.e. o reflexo social e a cascata) que, segundo o autor, seriam resultado direto da maior limitação do campo perceptual visual em comparação com o auditivo. Na crítica a esse trabalho, destaquei que o valor de sua análise pode ser mais bem estimado quando uma cisão menos categórica for feita entre conversas em LOs e conversas em LSs, uma vez que recursos visuais (bem como limitações do campo de visão) também fazem parte da experiência de ouvintes – embora normalmente, aspectos gestuais sejam desconsiderados em pesquisas com LOs.

Na análise do corpus da libras propriamente dito, um primeiro passo foi o de estabelecer critérios claros para definir onde se iniciam e onde se encerram os turnos

sinalizados. O conceito utilizado para essa definição foi o de *unidade gestual* (Kendon, 1972; Kita et al., 1997; McNeill, 1992), que, por sua vez, inclui as *fases do gesto* e as *seqüências gestuais* que elas compõem, desde o momento em que as mãos saem do repouso, até o momento em que elas retornam em direção a ele, parcial ou totalmente. Foi argumentado que fases do gesto como a preparação e a retração, amplamente ignoradas nos estudos de LOs, tendem a ser superestimadas nos estudos das LSs, dada a sua maior proeminência quando realizadas com mãos e braços. Nesse sentido, uma análise foi feita discutindo a necessidade de os estudos das LOs darem maior atenção à *preparação* dos gestos opcionais ou articulatórios que acompanham a “fala”, e de os estudos das LSs relevarem o peso da *retração parcial e total* nas sobreposições em conversas sinalizadas.

O segundo passo da análise foi o de ilustrar alguns tipos de sobreposição de fala que necessitam ser excluídos do escopo da presente investigação por não se mostrarem convencionalmente problemáticos aos participantes. Entre essas situações, foram identificados no corpus instâncias de sobreposições terminais (i.e. falas simultâneas bastante breves que são estruturalmente previstas devido ao processo contingente de projeção dos possíveis pontos de término da fala em curso) e de continuadores (i.e. falas minimalistas que emergem ao longo do turno em curso para ratificar a sua continuidade). A meu ver, a falta de discriminação de circunstâncias como essas, aliadas à falta de critérios explícitos para a segmentação da cadeia de fala sinalizada, enfraquece as afirmações de estudos como os de Baker (1977) e Coates e Sutton-Spence (2001), que vão de encontro com o modelo do um-de-cada-vez.

O terceiro e último passo da análise, então, envolveu a descrição dos principais recursos utilizados na conversa em libras para a resolução de sobreposições de fala problemáticas. Tal descrição foi situada em análises qualitativas sobre a dinâmica das ocorrências de sobreposição, desde a sua inicialização até a sua resolução. A análise mostrou, tal como teoricamente previsto, que as situações de sobreposição de fala na libras repercutem diretamente na produção dos participantes, que têm sua fala marcada por reciclagens, interrupções abruptas, aumentos de volume, desaceleração e aceleração, fechamento prolongado dos olhos e olhares divergentes, pedidos explícitos de atenção e itens lexicais para licenciamento de turnos extensos.

Com algumas exceções (i.e. o aumento de volume de sinais, interrupção brusca da sinalização, fechamento prolongado dos olhos, aceleração para a unidade seguinte,

itens lexicais de licenciamento de turnos), os recursos identificados neste estudo são os mesmos que Baker (1977) descreve em seu estudo com a ASL. A diferença entre ambos os estudos pode ser vista, entendo eu, como resultado do foco particular que cada autor adotou: ao passo que Baker se concentra nos “reguladores” empregados nas situações de “transição fluida” entre falantes (reguladores de início, continuação e troca de turnos), procurei me concentrar justamente nas situações em que essa fluidez é comprometida. A semelhança nos achados de ambos os estudos, por outro lado, pode ser explicada pelo fato de os pontos de transição entre falantes, nos quais tais reguladores emergem, estarem inevitavelmente propensos a sobreposições, ainda que esses recursos sejam desenhados precisamente para evitá-las.

Aliada à crítica teórica sobre o argumento de Baker (1977) e Coates e Sutton-Spence (2001), a análise empírica sobre as situações de sobreposição de fala na libras veio fortalecer a proposta da relevância do modelo do um-de-cada-vez na organização da conversação. A idéia de que os falantes surdos são capazes de produzir e observar a língua de sinais sem perda de compreensão, de Baker, ou de que os falantes surdos compartilhem continuamente a palavra na conversação, de Coates e Sutton-Spence, não encontra respaldo no comportamento dos participantes diante de situações de sobreposição de fala na conversa em libras. Em tais circunstâncias, os participantes surdos empregam recursos tanto para conquistar o direito pela palavra quanto para conceder o direito reivindicado por seu interlocutor, coordenando-os de maneira fina para encerrar a fala simultânea de uma maneira rápida e eficiente.

Assim como Baker (1977) e McIlvenny (1995) argumentaram, o olhar parece desempenhar um papel fundamental no gerenciamento da troca de turnos na conversa sinalizada. Diante de um olhar divergente, o participante surdo que ambicionava a palavra fazia uso de recursos tais como acenos e alongamento de sinais até que um olhar convergente fosse alcançado. Em nenhum momento a troca de turnos foi efetivada sem que o falante estivesse certo do contato visual do seu interlocutor. Além disso, a escolha criteriosa entre diferentes pedidos explícitos de atenção (i.e. HEI vs. FICAR) mostra que, quando ambiciona a palavra, o participante surdo trata como problemas interacionais distintos o fato de o interlocutor estar sinalizando com olhar convergente ou divergente. Assim, o argumento de Coates e Sutton-Spence (2001), de que o contato visual é desnecessário na interação sinalizada pelo fato de a visão periférica dos surdos ser bem mais apurada, novamente não encontra respaldo na conversação em libras.

Tampouco encontra respaldo a idéia das autoras de que esses pedidos explícitos de atenção, que segundo elas tornariam a interação “lenta e trabalhosa”, sejam incompatíveis com as conversas informais. Sacks (1992: 215-221), em sua aula intitulada “*Doing ‘being ordinary’*”, trata especificamente da questão de como ser cotidiano, comum, corriqueiro (*being ordinary*) exige, dos membros de uma sociedade, um grande investimento ou trabalho (*doing*). Sob esse prisma, o fato de os pedidos explícitos de atenção na conversa sinalizada tornarem-na “trabalhosa” simplesmente reflete o fato de os participantes na conversação estarem continuamente engajados num trabalho de fazer dela algo comum, rotineiro, mundano, tais recursos emergindo para solucionar um dos vários problemas endêmicos da interação que está continuamente colocando esse senso de normalidade em cheque: as situações em que mais do que um fala ao mesmo tempo.¹⁵⁸

O fato de os pedidos explícitos de atenção ocorrerem não apenas diante do olhar divergente, mas também do olhar convergente, revela ainda dois aspectos relevantes ao argumento da presente pesquisa. Em primeiro lugar, o que é relevante para o gerenciamento da troca de turnos na libras não é apenas a questão do “acesso perceptual”, como sugerem as discussões de Baker, McIlvenny e Sutton-Spence. Esse acesso é um pré-requisito, ou *condição necessária* para a efetivação da troca de turnos, como de fato argumentam Baker e McIlvenny, mas não *condição suficiente*.

Uma reflexão sobre essa “condição suficiente” nos remete, então, ao segundo aspecto relevante para a presente tese: o fato de serem vários os sinais comportamentais de um falante que servem como indicador do seu direcionamento e nível atencional. Nas instâncias estudadas, por exemplo, o fato de o participante continuar sinalizando fluentemente, ou reter a suspensão independente de um sinal no ar, é um sinal de que a sua atenção à contribuição de seu interlocutor se encontra menos ou mais comprometida, independentemente de seu olhar estar voltado ao interlocutor. Na medida em que, como argumentado no capítulo 4, a orientação do corpo e da cabeça também se constituam em manifestações comportamentais da atenção, é possível que esses outros recursos também interfiram na avaliação contínua que os participantes

¹⁵⁸ Aproveitando uma analogia de Schegloff (2000: 52), a “lentidão” e o “trabalho” que o uso desses recursos acarreta não é menos fundamental do que a “lentidão” e o “trabalho” acarretados por duas pessoas que, caminhando numa mesma calçada em sentidos opostos, se vêem na necessidade de coordenar os seus movimentos a fim de não se chocarem, desviando-se do seu curso mais simples e direto para conseguir uma passagem “normal”.

devem fazer de seus interlocutores em relação a eles estarem ou não prontos para responderem plenamente a uma nova contribuição.

A presente análise é um primeiro passo no estabelecimento de estudos de conversação em libras sob a orientação teórica da AC. Foi uma preocupação deste estudo não simplesmente assumir as premissas teóricas da área, como faz McIlvenny, mas de procurar inicialmente corroborar a operação do modelo do um-de-cada-vez, de forma que outros aspectos que tomam tal modelo por pressuposto possam ser posteriormente investigados com maior fundamento. Ao revelar que as sobreposições de fala se apresentam como sistematicamente problemáticas aos participantes, os achados da presente análise oferecem novas evidências de que a coordenação da ação social na conversação, independentemente de a língua que serve como recurso ser oral ou de sinais, requer a atenção conjunta sobre uma mesma entidade a cada novo momento. O capítulo 7 a seguir irá aprofundar a investigação sobre como podemos definir essa “entidade” lingüística numa língua de modalidade gestual-visual como a libras.

7. Análise II: As unidades entoacionais na libras

7.1. Introdução

O capítulo anterior demonstrou que alguns achados fundamentais no campo da AC, relativos à organização geral da troca de turnos na conversação, não se aplicam apenas às LOs, mas também a LSs tal como a libras. Com base nessa análise, pode-se afirmar que os interlocutores surdos, tanto quanto os ouvintes, se orientam com relação ao princípio do *um-de-cada-vez* para gerenciar a sua participação. Tal orientação se revela no contínuo trabalho colaborativo que os participantes realizam a fim de solucionar as situações em que sobreposições de fala problemáticas emergem.

Demonstrada a pertinência do modelo do *um-de-cada-vez*, referente à alternância entre *turnos de fala*, o objetivo do presente capítulo, então, é o de dar mais um passo nesta investigação sobre a segmentação explorando a estruturação interna de cadeias de fala extensas em unidades gramaticais de diferentes níveis. A fim de viabilizar uma investigação dessa natureza em uma língua sobre a qual nosso conhecimento se mostra ainda incipiente, a estratégia de análise elaborada envolveu duas abordagens distintas: uma primeira, focada nas UEs mínimas do discurso; e uma segunda, focada em práticas conversacionais complexas, estruturadas a partir de conjuntos dessas unidades mínimas.

No que concerne à primeira abordagem, a presente investigação sobre as UEs se beneficiou de estudos da lingüística funcional e cognitiva, bem como de estudos da gestualidade. Pesquisas sobre a forma e função das UEs nesses campos têm levantado alguns critérios formais potencialmente universais que operam na delimitação dessas unidades, além de destacar a sua função fundamental de salientação de informações. Complementam essa reflexão os estudos que revelam o papel de vários tipos de gestualidade no estabelecimento da atenção conjunta na interação. Utilizando esses conhecimentos como ponto de referência, o olhar sobre a cadeia de fala sinalizada pode demonstrar a relevância tanto de critérios de delimitação das UEs quanto da gestualidade manual e não-manual empregada nos processos de coordenação atencional.

No que concerne à segunda abordagem, a presente investigação sobre as UEs se beneficiou dos estudos da AC sobre práticas complexas que são metodicamente empregadas pelos participantes na organização da fala-na-interação. Essas práticas – *atividades* ou *condutas estruturadas*, nos termos de Schegloff (2002: 290) – como, por

exemplo, as narrativas, se desenvolvem como seqüências ordenadas de pequenas ações que, em conjunto, compõem ações maiores, estabelecendo uma trajetória que envolve um início e um término reconhecíveis. Esse complexo estruturado exige dos participantes uma orientação ao todo (e.g. porque ele me contou essa narrativa?) que, para ser adequadamente realizada, exige também orientação a cada uma das partes que o compõem (e.g. qual a contribuição dessa unidade para a narrativa de modo geral?). O fato de várias dessas condutas estruturadas se manifestarem em várias línguas permite, então, que a partir da identificação de suas partes, descritas na literatura da AC com base no estudo de LOs, possamos buscar as suas contrapartidas formais (i.e. os recursos por meio das quais tais partes tornam-se identificáveis) numa língua ainda pouco conhecida como a libras.

O capítulo se inicia com uma exposição do conhecimento que servirá de base para as duas abordagens dos dados acima descritas. Em seguida, serão introduzidos recortes metodológicos nos dados a fim de viabilizar esses dois olhares distintos sobre a conversação em libras. O primeiro recorte se refere à identificação de instâncias de turnos simples, possivelmente formados por uma única unidade, e o segundo, à identificação de instâncias de turnos complexos formados por listas e contrastes. Esse procedimento irá permitir que a análise dos dados propriamente dita caminhe dos casos menos problemáticos (turnos simples) para casos com um grau maior de dificuldade (turnos maiores porém estruturados).

7.2. Procedimentos de análise

Como argumentado acima, com o intuito de chegar às unidades gramaticais da libras, a presente investigação envolveu a combinação de duas abordagens: uma com enfoque nas UEs; e outra com enfoque em práticas estruturadas. No que concerne à primeira via, o presente estudo se utilizou do conhecimento acumulado em trabalhos tanto no âmbito da lingüística funcional e cognitiva quanto dos estudos da gestualidade. Esse conhecimento serviu como ponto de referência das análises de duas maneiras: pelo seu enfoque funcional, que envolve a questão do gerenciamento da atenção na interação; e por seu enfoque formal, que envolve o levantamento de critérios formais para a identificação de UEs, bem como de gestos empregados pelos interlocutores nos processos de coordenação atencional entre indivíduos.

Retomando brevemente a discussão do capítulo 4, estudos translingüísticos têm elencado ao longo dos anos uma série de critérios potencialmente universais para a segmentação do discurso em UEs: a pausa; o padrão de aceleração-desaceleração de unidades; os contornos entoacionais coesos e as redefinições de pitch. Ainda que todos esses diferentes recursos contribuam para o processo, alguns se constituem em critérios mais fortes do que outros, e a combinação de vários deles numa única unidade também é comum.

As pausas são os recursos mais salientes para a delimitação de UEs, embora elas possam ocorrer dentro de unidades, e não apenas nas suas fronteiras. O padrão rítmico de aceleração-desaceleração é outro recurso importante, manifestando-se na forma de reduções fonéticas nas sílabas fracas iniciais da unidade e alongamentos nas sílabas finais. Por fim, os contornos entoacionais, acompanhados das redefinições de pitch nas fronteiras das unidades, possivelmente são os critérios mais investigados pela maioria dos pesquisadores. Tendo em vista a universalidade em potencial desses critérios, a sua pertinência para a identificação de UEs na produção sinalizada poderá ser verificada ao longo da análise.

No que se refere à função das UEs, algumas metáforas utilizadas por Langacker (2001) se mostram interessantes para uma reflexão sobre a natureza dessas unidades na libras. O autor argumenta que a função das UEs seja a de realizar *enquadramentos atencionais* ao longo da cadeia de fala, possibilitando a coordenação de ambos os interlocutores sobre uma mesma entidade a cada novo momento. Além disso, concentrando-se sobre a dimensão entoacional da prosódia, o autor propõe que os contornos entoacionais devam ser entendidos como *gestos atencionais coesos*, tendo em vista a sua função de indicação dos limites de *janelas ou quadros de atenção*. Tal perspectiva é reforçada pelos estudos da gestualidade, que têm demonstrado o papel que desempenham vários tipos de gestos (i.e. o apontamento, o olhar, a cabeça, o tronco, bem como as suas inter-relações possivelmente hierárquicas) na salientação de informações e no estabelecimento da atenção conjunta na interação.

Essa relação entre prosódia/gestualidade, de um lado, e o gerenciamento da atenção, de outro, nos permite vislumbrar de que modo uma língua visual-gestual simbolizaria, ao menos no nível do discurso, essa função crucial das línguas humanas que é a salientação. Enquanto a análise de UEs como “janelas de atenção” nas LOs envolve um claro exercício de abstração, dada a dificuldade de mapearmos padrões

sonoros em imagens, em LSs como a libras, que dependem fundamentalmente da exploração do gesto e do espaço, é possível imaginar que a construção dessas “janelas” seja feita de maneira altamente icônica. Por exemplo, tendo em vista que o léxico das LSs é fundamentalmente produzido pelas mãos no espaço, “gestos atencionais” podem direcionar o interlocutor a diferentes entidades lingüísticas (i.e. sinais e frases) de maneira similar à que nós, humanos, por meio desses mesmos gestos, direcionamos a atenção de nossos interlocutores a entidades não-lingüísticas (i.e. objetos) no mundo externo ao discurso.

Tendo em vista as características formais das UEs e dos gestos atencionais nas LOs, então, foi feito um primeiro recorte nos dados transcritos por meio da seleção das trocas de turno mais simples, onde parecia mais evidente que o turno do participante fosse constituído por uma única unidade. Isso ocorria, principalmente, quando o falante fazia uso de práticas que selecionavam o interlocutor a tomar o turno por meio de uma frase simples, por exemplo, iniciando um reparo ou ainda a primeira parte de um par adjacente (e.g. convite/aceitação-rejeição, elogio/aceitação-rejeição). Nessa etapa de análise, assumiu-se que as características formais identificadas nas unidades isoladas, em especial no que diz respeito às suas fronteiras, seriam relevantes para a identificação de UEs em cadeias de fala mais extensas.

Num segundo momento, então, o olhar sobre unidades isoladas foi combinado com um olhar mais amplo, sobre cadeias de fala complexas (embora altamente estruturadas). Esse novo olhar motivou um segundo recorte nos dados transcritos: a seleção de práticas ou condutas estruturadas, como por exemplo as *listagens*, relativamente bem investigadas nas LOs, inclusive em línguas historicamente não-relacionadas (e.g. Jefferson, 1990, Lerner, 1994, e Schiffrin, 1994, para o inglês; Johnstone, 1983, para o árabe; Sánchez-Ayala, 2003, para um contraste entre o inglês e o espanhol; Selting, 2003, para o alemão; Tao, 1996, para o mandarim, entre outros). Vários desses estudos mostram que a prosódia desempenha um papel fundamental na indicação das partes componentes das listas, de modo que um contraste entre o modo como essa prática é estruturada nas LOs e o modo como ela é estruturada na libras poderia trazer novos indícios sobre as características que a prosódia assume nas LSs.

De especial interesse para esse segundo momento de análise é o trabalho de Selting (2003), que faz uma síntese da literatura sobre listas, buscando conciliar os estudos laboratoriais da prosódia com os estudos da AC sobre esse mesmo tema. A

autora mostra que, no que concerne à abordagem sociológica, os trabalhos sobre listas têm se concentrado, por meio da análise de conversas espontâneas, no aspecto sequencial dessas práticas. Eles apontam a existência de uma estrutura tripartida, formada por uma unidade que indica o início da lista, o detalhamento da lista em si, e uma unidade que indica seu fechamento. Diferentemente, os estudos laboratoriais da prosódia, por meio de experimentações e análises acústicas, têm se concentrado na distinção entre diferentes tipos de listagem (i.e. aquelas com um número fechado de itens e aquelas sem um número de itens pré-definido) de acordo com os diferentes padrões entoacionais empregados. Conciliando as duas propostas, Selting buscou investigar de que maneira os recursos prosódicos – além dos recursos estruturais e sequenciais levantados na AC – contribuíam interacionalmente para a realização dessa prática em conversas espontâneas.

Em sua análise, Selting argumenta que as listas, além de serem por si só constituídas de três partes, caracterizam-se como *práticas encaixadas*; isto é, elas são a parte do meio de uma atividade maior, que envolve um componente disparador no início (uma colocação genérica a ser detalhada), o detalhamento dessa colocação (através da lista *per se*), e um fechamento gestáltico ao final (que retoma a relevância da lista no contexto interacional). Embora a última parte dessa estrutura não seja imprescindível, Selting mostra uma forte preferência dos falantes por preservá-la. No que diz respeito à prosódia, Selting mostra que ela funciona, isoladamente ou junto com outros recursos, na indicação de aspectos como: a) o início da prática de listagem; b) o tipo de lista em curso (i.e. fechada ou aberta); e c) o estatuto da unidade no corpo da lista (i.e. se ele é um item listado ou uma digressão; e se é um item não-final ou final). Por fim, ela assinala que o forte paralelismo sintático e prosódico (por meio da recorrência de contornos prosódicos em cada item listado) confere uma forte unidade à atividade e reforça a orientação dos participantes às listas como uma prática social coesa. Tal orientação se revela na troca fluida e coordenada de turnos que ocorre de maneira sistemática na fronteira final dessas práticas.

O trabalho de Sánchez-Ayala (2003) também apresenta uma relevância especial para a presente análise por duas de suas características principais: primeiro, a de estar focado na comparação entre a realização de listas no inglês e no espanhol, ao invés de simplesmente descrever o que ocorre em uma ou outra língua; segundo, a de buscar

complementar a motivação interacional das listas com uma análise sobre a motivação cognitiva para essa mesma prática.

No que se refere ao contraste entre inglês e espanhol, o autor mostra que, em ambas as línguas, as listas apresentam padrões formais e funcionais similares. Entre as similaridades formais estão os padrões prosódicos estilizados e paralelísticos, a coordenação sintática e a opção de utilizar um número restrito de itens lexicais e expressões (e.g. “tal y cuál”, “and all that stuff”), os chamados *finalizadores*, que completam a lista com uma generalização. A única diferença identificada pelo autor está no fato de o espanhol apresentar três padrões prosódicos recorrentes (relativos aos tons terminais que recorrem a cada nova unidade entoacional), ao passo que, no inglês, apenas um padrão recorrente é observado, com muitos casos revelando padrões paralelísticos inéditos que respondem de maneira idiossincrática às situações imediatas onde cada lista emerge.

No que se refere às motivações das listagens, Sánchez-Ayala argumenta, primeiramente, que tais práticas sejam utilizadas para a realização de duas tarefas interacionais: o estabelecimento de uma base de conhecimento comum entre os interlocutores, antes de dar continuidade a uma atividade maior (*framing lists*); e o oferecimento de evidências para um argumento por meio de uma descrição detalhada (*demonstrative lists*). Essa diferença funcional se reflete no fato de os receptores de uma lista em curso, no caso das listas “*framing*”, reagirem sistematicamente *ao longo de* seu curso com sinais de acompanhamento, enquanto, no caso das listas “*demonstrativas*”, a reação aparece apenas *ao término da* prática. Do ponto de vista cognitivo, o autor argumenta que o formato gramatical das listas (i.e. forte paralelismo sintático e prosódico, recorrência parcial e total de itens lexicais e expressões, além da presença de pausas e completadores (*fillers*)) sugere que tais práticas sejam “*adaptações ao planejamento sobrecarregado de unidades entoacionais tematicamente equivalentes no discurso espontâneo*” (p. 338).¹⁵⁹

¹⁵⁹ Esse argumento remete à análise de Ong (1982) sobre como as culturas orais resolvem, pela exploração de recursos lingüísticos, problemas cognitivos inerentes à produção e transmissão de longos textos. Recuperando a análise sobre os poemas Homéricos de Milman Parry – que demonstrou como esses poemas podiam ser passados oralmente de geração para geração –, Ong afirma que a exploração de métodos orais de composição, tais como a recorrência parcial ou total de itens lexicais e expressões, bem como o paralelismo sintático e prosódico, atuavam como “*fórmulas*” que facilitavam a memorização do texto e ofereciam a flexibilidade necessária aos poetas para a improvisação intrínseca às performances.

Com esse conhecimento sobre a prática de listagem em LOs, pude dispor de uma base de comparação para a análise das listas na libras selecionadas nesse novo recorte dos dados. Entre os questionamentos pertinentes a essa análise, destacam-se as seguintes questões: i) a lista na libras apresenta a mesma estrutura tripartida que nas LOs? ii) a lista na libras se apresenta como uma prática encaixada numa seqüência interacional que a contextualiza? e iii) o papel que a lista na libras desempenha traz alguma relação com a contextualização de informações ou a demonstração de um argumento? Se sim, podemos então passar às questões que mais diretamente interessam à presente investigação: iv) com quais recursos formais um participante mostra ao outro que uma prática de detalhamento (i.e. uma lista) foi iniciada, está em curso, e foi encerrada? v) com quais recursos ele mostra o estatuto dos diferentes componentes internos à lista (i.e. como itens listados ou como digressão)? e vi) que tipo de recorrências e paralelismos na libras conferem à prática uma maior coesão? Tal análise, além de permitir a verificação dos achados alcançados a partir da análise do primeiro recorte nos dados, permitiu também a identificação de novos recursos prosódicos relevantes para a questão da segmentação do discurso sinalizado em unidades gramaticais.

O segundo recorte sobre os dados envolveu ainda a seleção de uma outra prática ou conduta estruturada: o contraste. Mais uma vez, neste caso, tive a oportunidade de utilizar como referência para a análise um estudo sobre esse mesmo fenômeno realizado numa língua oral, o inglês (Atkinson, 1984).¹⁶⁰ Embora o trabalho de Atkinson não tome por base a conversação, mas sim discursos políticos, os paralelos entre o que o autor descreve e as práticas com as quais me defrontei no corpus são marcantes, de modo que esse estudo também se mostrou relevante como base de comparação para a análise da libras.

Nos termos de Hutchins (2005), pode-se argumentar que certos gêneros de fala (e.g. poemas épicos) e certas práticas interacionais (e.g. listas) constituam produtos culturais que emergem, pelo menos em parte, como adaptação às nossas limitações cognitivas. Tal adaptação se dá por meio da exploração da recorrência lexical, sintática e prosódica como âncoras materiais (i.e. formadas por matéria gestual e auditiva, no caso das LOs, e gestual e visual, no caso das LSs) para a produção e compreensão de trechos de discurso com maior grau de complexidade.

¹⁶⁰ Na verdade, a análise do contraste foi inicialmente feita sem referência ao estudo de Atkinson (1984), que só foi descoberto posteriormente. Apesar disso, essa leitura permitiu-me identificar outras ocorrências no corpus que a princípio eu não havia enquadrado como instâncias de contraste, possibilitando uma visão mais ampla sobre essa prática e, conseqüentemente, um maior aprofundamento da análise.

Atkinson aponta que o método do contraste, ou da “antítese verbal”, tem uma longa história nos estudos da retórica e oratória clássicas, constituindo-se numa das principais “figuras de linguagem artísticas” desse gênero. O problema que motivou o autor a identificar e analisar essa prática foi proveniente dos próprios dados do autor: a necessidade de explicação de como o público, num discurso político, era capaz de oferecer respostas (i.e. aplausos) de maneira orquestrada. Ao analisar as falas que precediam esses aplausos, então, Atkinson identifica os contrastes como uma das práticas mais eficazes na eliciação de respostas afiliativas do auditório.

Do ponto de vista formal, contrastes são construções que, por meio da realização de uma primeira parte, permitem a projeção de uma segunda parte pelo receptor. Para possibilitar a identificação tanto do contraste em si quanto do seu momento de possível completude, os oradores utilizam recursos tais como: repetir itens lexicais ou expressões da primeira parte do contraste na segunda, com pequenas modificações; manter a extensão da duração de cada uma das partes relativamente igual; e aumentar o volume e o ritmo da fala na aproximação do término da segunda parte. Para Atkinson, o emprego desses últimos recursos, de natureza prosódica, deve estar relacionado à questão do gerenciamento da atenção, com o maior esforço investido na fala (por meio da manutenção de um novo nível de amplitude e uma maior ênfase rítmica) sinalizando algo de especial importância, frente ao qual o público deve prestar maior atenção (p. 400).

Do ponto de vista funcional, Atkinson argumenta haver evidências de que os contrastes sejam práticas recorrentes em um grande número de situações nas quais a persuasão e o convencimento do(s) receptor(es) constitua o propósito central do falante. Isso explicaria a grande relevância dessa prática no debate político, frequentemente marcado pelo caráter polarizado de oposição *vs.* situação, e vice-versa. Sobre as possíveis relações entre o uso dessas práticas no discurso político e na conversação, então, o autor argumenta haver evidências de que os recursos empregados na sua construção, em especial os prosódicos e gestuais, difiram em alguns aspectos de um gênero de fala para outro, o que sugere cautela na extrapolação dos achados. Apesar disso, o autor destaca que, “na medida em que [os recursos] demonstrem relevância para resolver o problema de atrair, manter e/ou ampliar a atenção de membros do auditório, pode-se um dia vir a demonstrar que as atividades não-vocais e prosódicas em contextos

públicos se assemelham muito (talvez de uma maneira exagerada) à sua operação em interações cotidianas de menor escala” (p. 408).

Assim como procedido em relação à prática de listagem, então, foi possível utilizar o conhecimento sobre o contraste no inglês como base de comparação para a análise dos contrastes na libras identificados neste segundo recorte nos dados. Entre os questionamentos pertinentes a essa análise, destacam-se as seguintes questões: i) há alguma correlação entre a recorrência de contrastes na conversa em libras e o propósito de persuasão e convencimento do interlocutor? ii) as conversas em que os contrastes são predominantes apresentam um caráter polarizado, similar ao dos discursos políticos?; e iii) os contrastes na interação conversacional são igualmente eficazes em eliciarem uma pronta resposta do(s) interlocutor(es)? Se sim, podemos então passar às questões que mais diretamente interessam à presente investigação: iv) com quais recursos formais um participante surdo mostra ao outro que uma prática de contraste foi iniciada, permitindo a antecipação da ocorrência de uma segunda parte? v) que desenho específico do contraste na libras permite que o seu término possível seja identificado? e vi) que tipo de similaridade e/ou diferença formal entre cada uma das partes vem reforçar o contraste como um todo coeso? Mais uma vez, aqui, a resposta a essas perguntas permitiu tanto a verificação de achados prévios quanto a identificação de novos recursos prosódicos relevantes para a segmentação do discurso em unidades gramaticais.

Diante dos achados obtidos a partir das análises sobre as instâncias de turnos simples, bem como de condutas estruturadas tais como listas e contrastes, uma base de conhecimento inicial sobre a prosódia da libras pôde ser formada. Isso permitirá que, num momento subsequente, o discurso em libras como um todo possa ser abordado com maior segurança e fundamentação. Essa extrapolação da investigação para além desses dados “controlados”, porém, é uma etapa do estudo que, embora necessária, não entrou no escopo desta tese, ficando reservada para pesquisas futuras.

7.3. Instâncias de turnos simples

Na análise de turnos simples, isto é, possivelmente formados por uma única unidade, procurei observar padrões recorrentes na sinalização que pudessem despontar como possíveis delimitadores de UEs, dando especial atenção às modulações dos sinais manuais. Além disso, procurei levantar hipóteses sobre aspectos manuais e não-manuais

da sinalização que poderiam indicar algo como acento (i.e. proeminência) de elementos no âmbito da UE. Nessa primeira etapa de investigação, alguns padrões relevantes já puderam ser identificados, de modo que as análises serviram como ponto de apoio para a tarefa subsequente de segmentação de turnos mais complexos, formados por listas e contrastes.

Na investigação de onze instâncias de turnos simples, o *padrão de aceleração-desaceleração* destacou-se como um dos critérios mais representativos, razão pela qual esse critério, distribuído nas seções sobre *alongamento final* e *aceleração inicial*, aparece como o primeiro da discussão sobre as UEs na libras. Em seguida, passo para a análise do que, aproveitando a metáfora de Langacker (2001) sobre a entoação, optei por chamar de *gestos atencionais coesos*. Como um último aspecto relevante para a delimitação de UEs, então, passo à discussão das *piscadas de olhos*. No que diz respeito a possíveis marcas de *acentuação*, irei discutir uma única instância que, tanto pela forma como pela função no discurso, me pareceu ser um forte candidato a acento na libras.

7.3.1. Alongamento final

O maior número de hipóteses levantadas nesta etapa da análise esteve relacionado ao padrão de aceleração-desaceleração dos sinais e, mais particularmente, à ocorrência de alongamentos finais, que pôde ser identificada em quase todas as instâncias de turnos simples selecionadas. Não por acaso, Chafe (1994), entre outros estudos, aponta esse padrão como o mais recorrente, destacando-se como um dos critérios mais robustos para a identificação de UEs.

Nas LOs, o alongamento final caracteriza-se, de maneira geral, como um alongamento da última sílaba acentuada da frase. Na libras, a partir das observações das instâncias de turnos simples, proponho que sejam consideradas manifestações de alongamento final pelo menos três tipos de modulações presentes em sinais de posição final. A primeira delas é a *manutenção da suspensão pós-golpe* em sinais cuja fase expressiva é realizada por meio de golpe, ou ainda a *manutenção da suspensão independente*, em sinais que não possuem golpe.

Dois exemplos do corpus aparecem na figura 36 abaixo. O primeiro exemplo, na parte de cima da figura, mostra o sinal alongado SER-APROVADO, que, no contexto, aparece como único sinal da frase (“Foi aprovado?”). As imagens referem-se aos três

momentos da fase expressiva desse sinal: i) a suspensão pré-golpe, com duração de dois frames; ii) o golpe, com duração de seis frames; e iii) a suspensão pós-golpe, com duração de seis frames.

O segundo exemplo, na parte de baixo da figura, mostra o sinal alongado “A”, que aparece no contexto como a última letra da palavra soletrada manualmente “física”.¹⁶¹ As imagens mostram os três últimos sinais soletrados, “T”, “C” e “A”, os três formados por suspensões independentes. Nota-se, então, que enquanto “T” e “C” apresentam três frames de suspensão cada, o sinal ‘A’, o último da frase, é alongado por cinco frames de suspensão.

O segundo tipo de modulação dos sinais em posição frasal final é a *reiteração dos movimentos repetitivos internos ao golpe*, quando ele está presente.¹⁶² Um exemplo que aparece no corpus é o sinal O-QUE, que em sua <[forma de menção](#)> é realizado com um golpe de dois movimentos repetitivos e que, no corpus, ao final da frase EXPLICAR O-QUE (“Explicar o que?”), aparece <[alongado com quatro movimentos repetitivos](#)>.

¹⁶¹ A soletração manual é um recurso da libras por meio do qual os surdos soletram palavras das LOs utilizando um alfabeto manual. Esse recurso é utilizado principalmente para referir-se a nomes próprios (de pessoas, ruas, instituições, entre outras), mas pode também ser utilizado para outros fins específicos.

No contexto em questão, o falante quis comunicar à sua interlocutora o sentido de um sinal desconhecido por ela, empregando para isso uma expressão equivalente do português (a palavra “física”). Alguns sinais do alfabeto manual, como o ‘h’, ‘j’, ‘k’, ‘x’ e ‘z’ apresentam golpe, mas a grande maioria apresenta apenas uma suspensão independente, como é o caso das três letras apresentadas na parte inferior da figura 36.

Embora, viesados pelo português, possamos pensar que a soletração de “física” deva fazer referência às letras de uma palavra, do ponto de vista da libras cada letra manual soletrada pode ser entendida como um sinal e a palavra ou expressão soletrada, uma frase – embora seja plausível especular que esse tipo de frase deva apresentar características formais e funcionais peculiares quando comparada às frases mais convencionais da libras.

¹⁶² De acordo com o trabalho de Kita et al. (1997), quando um dado gesto apresenta movimentos repetitivos (e.g. a representação gestual de bater em algo com um martelo), toda a duração da repetição é contada como uma única fase repetitiva (i.e. um golpe formado por movimentos repetitivos, ao invés de cada movimento ser considerado um golpe separado). Relembrando os critérios para segmentação dos gestos, devem ser considerados nessa questão não apenas as mudanças de direção do movimento, mas também do perfil de velocidade. Vários dos sinais da libras, como o sinal O-QUE discutido nesta seção, apresentam essa característica repetitiva do golpe.



Figura 36. Alongamento da suspensão pós-golpe e da suspensão independente de sinais em posição frasal final.

Além de meu conhecimento intuitivo da libras sobre a forma de menção desse sinal, tal análise pôde ser corroborada em mais de uma ocorrência pelos próprios dados do corpus, pois um mesmo sinal, quando aparecia em contexto frasal inicial ou medial, apresentava um golpe com número de repetições igual ou inferior à forma de menção, porém, quando em contexto frasal final, apresentava o golpe com repetições adicionais.

É o caso, por exemplo, do sinal OBRIGAR, que, em sua <forma de menção>, é realizado com duas repetições. Quando esse sinal aparece num contexto medial identificado no corpus, na frase <VOCÊ OBRIGAR ESS@-DOIS IGUAL> (“Você vai obrigar os dois a serem iguais?!”), o sinal aparece tal qual na forma de menção; contudo, quando aparece num contexto final, na frase <VOCÊ OBRIGAR> (“Você vai obrigar?”), o mesmo sinal revela um golpe de três movimentos repetitivos.

Além disso, tanto o exemplo do sinal OBRIGAR quanto o do sinal O-QUE permitem observar outros padrões interessantes relacionados aos sinais formados por golpes com movimentos repetitivos. Primeiramente, quando esses sinais são alongados ao final da frase, predomina nos dados uma reiteração *parcial* do golpe, de modo que sinais envolvendo golpe com dois movimentos repetitivos frequentemente acabam realizados com três. Em segundo lugar, quando se observa a dinâmica dessas

reiteraões, nota-se que há uma diminuição progressiva na extensão do deslocamento do movimento em cada repetição, da primeira até a última, produzindo um efeito nítido de decaimento. Essas observações sobre a dinâmica dos sinais em posição frasal final, a meu ver, vão de encontro com a proposta de Nespor e Sandler (1999), de que os sinais em posição final nas *frases fonológicas*¹⁶³ envolvam uma “reduplicação” relacionada à marcação de proeminência, e não ao fenômeno do alongamento final.

O terceiro e último tipo de modulação observado nos sinais em posição frasal final é a *transformação de um sinal cuja fase expressiva é formada por uma suspensão independente em um sinal do tipo alternativo, formado por um golpe*. Um exemplo que aparece no corpus é o sinal VOCÊ, que em sua <[forma de menção](#)> é realizado por meio de um gesto simples de apontamento que culmina numa suspensão independente, e que, no corpus, em posição final na frase <[PROFESSOR EXPLICAR VOCÊ](#)> (“Você ia explicar para o professor?”), é transformado num sinal com golpe repetitivo, similar ao sinal O-QUE ou OBRIGAR.

A identificação do alongamento final na libras vem corroborar os achados de várias pesquisas da ASL (Baker, 1977; Coulter, 1993; Liddell, 1978; Wilbur e Nolen, 1986, entre outros), o que, tendo em vista a sua ampla constatação também nas LOs, qualifica esse fenômeno como possivelmente universal. Como apontam Myers e Hansen (2006), a universalidade do alongamento final tem levado alguns pesquisadores a questionar a possível motivação desse fenômeno. Segundo os autores, a hipótese mais plausível, encontrada em trabalhos como os de Klatt (1976) e Edwards et al (1991), é a de que o efeito do alongamento fonético final deva estar relacionado a uma desaceleração da articulação em decorrência da antecipação que os falantes fazem de uma pausa subsequente, o que exige uma interrupção no movimento dos articuladores. Como podemos observar nas provas de velocidade no atletismo, nas quais os atletas diminuem gradualmente a sua velocidade após o término da prova para evitar contusões, parece ser mais fácil fisiologicamente fazer a transição entre uma condição de movimentação e uma condição estática de maneira gradativa.

¹⁶³ Embora essa crítica esteja em certo sentido comprometida pelo fato de as autoras fazerem tal proposta sob uma teoria prosódica específica e não compartilhada por este estudo, os dados trazidos em seu trabalho me fazem crer que o que elas chamam de *frases fonológicas* corresponda, grosso modo, às UEs tal como tem sido aqui descritas.

Uma evidência para essa hipótese está na observação de que, nas ocorrências de alongamento final, quanto mais próximo do final, mais um dado segmento será estendido; gradiente essa que já foi atestada em línguas diversas como o sueco (Lindblom et al. 1981), o árabe (de Jong and Zawaydeh 1999) e o inglês (Silverman and Pierrehumbert 1990; Turk 1999). Se considerarmos o golpe como um segmento na cadeia de fala sinalizada – independentemente do estatuto gramatical que deva ser a ele atribuído – a análise da libras pode servir como mais uma evidência desse caráter progressivo da cessação do movimento a medida que se aproxima o término de uma UE.

A meu ver, tal explicação vem reforçar a relevância da prosódia para a segmentação do discurso em unidades gramaticais, uma vez que é a operação mental de planejamento e projeção, em tempo real, de uma unidade relevante cognitivamente (i.e. a previsão de um ponto de início e um ponto de chegada na fala imediata) que, em última instância, exige essa adaptação motora de nosso sistema fisiológico.

7.3.2. Aceleração inicial

Nas LOs, a aceleração inicial revela-se pela redução fonética das sílabas fracas que antecedem a primeira sílaba acentuada da frase, processo freqüentemente referido como *anacruse*. Na libras, a partir das observações das instâncias de turnos simples, quatro tipos de fenômenos em posição frasal inicial emergem como candidatos a manifestações de aceleração inicial.

O primeiro deles é *a sobreposição de fases do gesto*, que ocorre quando uma mão inicia a preparação de um novo sinal enquanto a outra ainda realiza a fase expressiva de um sinal anterior.¹⁶⁴ O exemplo da figura 37, relativo à frase <[VOCÊ TER INTÉRPRETE](#)> (“Você tinha intérprete?”), ilustra esse fenômeno, com as fases de preparação de dois sinais, TER e INTÉRPRETE, sobrepondo-se à fase expressiva dos

¹⁶⁴ Como discutido na seção 2.2, Battison (1978) argumenta que, apesar de as LSs possuírem dois articuladores ativos idênticos (duas mãos-braços), há restrições no uso simultâneo desses articuladores: as chamadas *condição de dominância* e *condição de simetria*. A sua análise, contudo, assim como grande parte das análises fonológicas no campo das LSs, toma por objeto de estudo apenas a fase expressiva de sinais isolados, ao invés de também levarem em consideração, como propõem Kita et al. (1997), as fases de preparação e retração.

Assim, quando seguimos a orientação de Kita et al. e observamos a relação entre essas diferentes fases do gesto na fala sinalizada espontânea, notamos uma possibilidade adicional de interação entre os dois articuladores principais da libras. Tal fato pode ter implicações para uma teoria fonológica/prosódica sobre as LSs (e.g. proposta de Sandler, 1999, sobre cliticização na discussão da nota de rodapé 19).

sinais que os precedem na frase, respectivamente, VOCÊ e TER. Desse modo, o sinal VOCÊ, realizado com a mão esquerda, é acompanhado da preparação do sinal TER (imagens a e b), e o sinal TER, por sua vez, é acompanhado da preparação do sinal INTÉRPRETE (imagem c).¹⁶⁵





			
a. Neste frame, a mão esquerda está no final da preparação de VOCÊ, quando a mão direita dá início à preparação de TER, saindo do repouso	b. Quando VOCÊ alcança a suspensão independente, a sua fase expressiva, a mão direita já está no meio da preparação de TER	c. Enquanto a mão direita finaliza o golpe de TER, a mão esquerda já antecipa a preparação do sinal INTÉRPRETE	d. A frase se encerra com o sinal INTÉRPRETE, alongado por meio do aumento das repetições internas ao golpe

Figura 37. Sobreposição de fases do gesto de sinais em posição frasal inicial.

A segunda modulação observada em sinais de posição frasal inicial é *a redução do número de repetições em golpes formados por movimentos repetitivos*. Trata-se de uma modulação exatamente oposta ao alongamento que ocorre em sinais com golpes repetitivos. Um exemplo que aparece no corpus é o do sinal CONHECER, que em sua <forma de menção> é realizado com um golpe de dois breves movimentos tocando o queixo, e que, no corpus, no início da frase <CONHECER FÍSICA> (“Sabe física?”), aparece com apenas um movimento sem repetição.

A terceira modulação observada nos sinais em posição frasal inicial é *a retenção bastante breve, por um ou dois frames – (menos de 0,1 segundo) – da suspensão independente de um sinal sem golpe*, praticamente reduzindo o sinal à realização plena

¹⁶⁵ Nespor e Sandler (1999: 19-20) se referem a essa antecipação da mão passiva relativa ao sinal subsequente como uma evidência da existência de certos domínios prosódicos, tal como a *frase fonológica*. Para as autoras, esse “espraiamento da mão não-dominante” de sinais feitos com as duas mãos, como ocorre com INTÉRPRETE, encontra como barreira justamente a fronteira desse domínio.

Cabe assinalar que um tipo de sobreposição de fases similar foi também identificado em contextos frasais finais. Como pode ser visto na frase EXPLICAR O-QUE (ver discussão sobre alongamento final acima), o sinal final O-QUE é parcialmente integrado na fase expressiva do sinal anterior, EXPLICAR. Para isso, a falante rompe com a simetria no golpe de EXPLICAR, iniciando a preparação do sinal O-QUE com a mão direita enquanto a mão esquerda finaliza o golpe do sinal anterior. Sandler (1999: 18-19; cf. nota de rodapé 19) analisa um exemplo similar a esse como uma ocorrência de cliticização de pronomes fracos a itens lexicais plenos.

de sua fase de preparação. Assim, a mão se movimenta na fase de preparação até atingir o ponto de articulação e a configuração de mão da fase expressiva, mas, ao invés de permanecer nessa suspensão independente, passa imediatamente para a preparação do sinal seguinte. Um exemplo que aparece no corpus é o do sinal VOCÊ, que no início da frase VOCÊ OBRIGAR ESS@-DOIS IGUAL (ver discussão sobre alongamento final acima) é realizado com apenas um frame de suspensão independente.

A quarta e última modulação dos sinais no início de frases é a *assimilação da configuração de mão do sinal subsequente pelo sinal inicial*, em sua fase expressiva ou até mesmo em sua fase de preparação.¹⁶⁶ Quando essa assimilação ocorre, o ponto de articulação e o movimento da fase expressiva do sinal inicial são preservados, mas a configuração de mão adquire as características do sinal subsequente.

Um exemplo que aparece no corpus é a assimilação de configuração de mão do sinal OBRIGAR pelo sinal VOCÊ, na frase VOCÊ OBRIGAR (ver discussão sobre alongamento final acima). Como mostra a figura 38 abaixo, até certo ponto da fase de preparação do sinal VOCÊ (imagens a e b), a falante ainda preserva a configuração de mão em “G1”, própria do sinal VOCÊ. Quando o sinal atinge a fase expressiva (imagem c), contudo, a configuração de mão em “A”, própria do sinal subsequente na frase, OBRIGAR (imagem d), é assimilada.^{167,168}







			
a. Início da preparação do sinal VOCÊ, com a mão direita, após o término da frase antecedente	b. Meio da preparação do sinal VOCÊ, com a configuração de mão própria do sinal ainda preservada	c. A configuração de mão de VOCÊ é assimilada pela configuração do sinal seguinte, OBRIGAR	d. O sinal OBRIGAR é realizado com sua configuração de mão característica no fim da frase

Figura 38. Assimilação da configuração de mão de sinais em posição frasal inicial.

¹⁶⁶ Utilizo o termo *assimilação* aqui num sentido não-técnico, reservando para pesquisas futuras a investigação sobre a possível relevância desse tipo de ocorrência para o fenômeno de “assimilação” – tal como tem sido descrito por fonólogos.


¹⁶⁷ Configuração de mão em “G1”:  e configuração de mão em “A”: .

¹⁶⁸ Para Sandler (1999: 23-25), esse tipo de assimilação é também interpretado como resultado da cliticização de pronomes fracos em itens lexicais plenos.

Três observações sobre essa última modulação se mostram relevantes. Em primeiro lugar, nota-se que a preparação da *localização* do sinal, que envolve a movimentação do(s) braço(s), e a preparação da *configuração de mão*, que é interna às mãos, embora freqüentemente associadas, são em última instância independentes e podem ocorrer de maneira dissociada. No caso em questão, a preparação interna das mãos, relativa ao sinal OBRIGAR, alcança o seu término mais rapidamente do que a preparação da localização desse mesmo sinal.

Em segundo lugar, o fato de um sinal como VOCÊ, cuja configuração de mão foi assimilada pela de outro sinal, ser ainda reconhecido como VOCÊ sugere que o movimento que prepara o sinal para o seu ponto de localização inicial seja mais saliente fonologicamente do que o movimento que prepara as mãos internamente para a sua configuração inicial.¹⁶⁹ Assim, temos uma pista importante para a segmentação de palavras numa frase: sempre que uma nova fase gestual de *preparação da localização* do sinal for iniciada, inicia-se o novo sinal. Embora observações do corpus mostrem que, de maneira geral, a preparação da localização é coordenada com a preparação da configuração das mãos em seu ponto inicial e final, uma equiparação das duas ou um favorecimento da última parece ser equivocado.¹⁷⁰

Uma terceira observação relevante sobre esse fenômeno da assimilação da configuração de mão no contexto da presente discussão é a de que, em si, ela não acarreta objetivamente nenhuma aceleração da fala – no sentido de que realizar dois sinais com assimilação não envolve necessariamente uma redução no tempo de sua produção, em contraposição à sinalização desses mesmos dois sinais sem assimilação. Apesar disso – assim como propus na análise da “aceleração para a unidade seguinte”

¹⁶⁹ Essa análise me remete a uma divertida conversa que tive um dia com um amigo surdo. Por algum motivo que não me recordo, comentei com ele que um surdo que tivesse as mãos amputadas estaria numa situação particularmente delicada, pois não conseguiria se comunicar. Meu amigo disse que não, que eu estava errado, e para demonstrar o seu ponto, manteve as mãos na configuração que mais se assemelhava à de mãos amputadas (a configuração em “S”, ) e produziu uma longa fala sem qualquer alteração nessa configuração. Exceto por um ou outro sinal, a sua fala de fato se mostrou surpreendentemente clara para mim.

¹⁷⁰ A questão da segmentação de palavras nas LSs é tão relevante e problemática quanto a segmentação de UEs ou turnos. Inicialmente, em discussões do projeto de formação de corpus do grupo ECS, havíamos estabelecido que essa segmentação seria feita inicialmente pela identificação de modificações nas configurações de mão dos sinais, pelo fato de, intuitivamente, tal traço nos parecer fonologicamente mais saliente. Essa solução não se sustenta, contudo, diante dos dados aqui apresentados, que sugerem que a preparação espacial das mãos para o local referente à fase expressiva do sinal seja mais relevante do que a preparação interna da configuração de mão.

na seção 6.4.3 –, há, a meu ver, um *efeito de aceleração* por meio da antecipação de características formais do sinal subsequente pelo sinal anterior, cuja forma se torna enfraquecida.

Por fim, cabe assinalar que o fato de reduções fonéticas ocorrerem em outras posições na frase além da posição inicial (cf. nota de rodapé 165) sugere que essas formas de modulação, antes de caracterizarem propriamente uma “aceleração inicial”, integrem, na verdade, processos fonológicos típicos de pontos de *disjunção mínima* na cadeia segmental, revelando a coesão interna de uma unidade (Local e Kelly, 1986: 186). A essas formas de modulação contrapõem-se as de caráter *disjuntivo*, tal como o alongamento final, que estabelecem uma fronteira entre uma cadeia segmental e outra. Tendo em vista que a porção inicial de uma unidade constitui-se como um ponto de juntura prototípico, reduções fonéticas dessa natureza devem se mostrar recorrentes nesse contexto, embora não estejam necessariamente vinculadas a ele.

7.3.3. Gestos atencionais coesos

Um outro elemento que se mostrou relevante para a questão da segmentação na análise dos turnos simples é o que optei por chamar de *gesto atencional coeso* – aproveitando o *insight* de Langacker a respeito da forma e função dos contornos entoacionais nas LOs. Nas instâncias de turnos simples, destaca-se como exemplo de gestos dessa natureza as inclinações de cabeça e expressões faciais que realizam os dois tipos básicos de interrogativas na libras (qu- e sim/não) – que podem ser vistas em sua forma mais prototípica, respectivamente, nas frases EXPLICAR O-QUE (cf. seção 7.3.1) e na frase VOCÊ OBRIGAR (cf. figura 38).

A relevância desse tipo de configuração gestual ficará mais evidente na análise das práticas de listagem e contraste.¹⁷¹ Neste momento, gostaria apenas de chamar a atenção para o fato de esses recursos atuarem como delimitadores de unidades gramaticais, funcionando como gestos suprasegmentais que permanecem relativamente estáveis durante a articulação do conteúdo segmental que está sendo enquadrado – similar ao que a entoação realiza nas LOs.

¹⁷¹ Na verdade, foi somente após a análise das práticas de listagem e contraste que pude identificar esse recurso, o que me fez retornar às instâncias de turnos simples e poder apontá-los. A razão é que tais gestos são mais evidentes quando inseridos num contexto de outras unidades produzidas pelo mesmo falante, e não em contextos de unidades isoladas, como é o caso desta seção.

A meu ver, é importante diferenciar essa proposta de compreensão sobre as expressões faciais nas LSs de visões tais como a de Nespor e Sandler (1999). Com base na análise de configurações do rosto e da cabeça (e.g. as interrogativas acima mencionadas), as autoras argumentam que esses sinais não-manuais devam ser vistos como análogos à entoação nas LOs. A diferença em relação à entoação seria a de que, ao passo que a entoação se caracteriza basicamente como uma *sucessão linear de variações de um único elemento* (i.e. a frequência percebida como pitch), as expressões faciais se caracterizam pela *estabilidade e simultaneidade de vários articuladores distintos* (i.e. boca, bochecha, sobrancelha, olhar, posição da cabeça) (p. 27).

Em minha proposta, porém, prefiro não equiparar os gestos atencionais a uma noção de “entoação sinalizada”, pelo menos neste momento. Primeiramente porque as configurações não-manuais que constituem tais gestos não existem apenas nas LSs, algumas delas podendo inclusive ser empregadas produtivamente junto com a fala nas LOs (como sugerem, e.g., análises nas linhas de Bolinger, 1983). Em segundo lugar, porque, olhando a prosódia sob uma perspectiva mais ampla, tanto do ponto de vista formal, como uma forma de gestualidade, quanto funcional, como um recurso de gerenciamento da atenção, parece-me mais plausível a idéia de que as LSs devam suprir a ausência do pitch por meio de recursos gestuais mais icônicos: o olhar, a orientação da cabeça e do tronco, o apontamento, a espacialização dos sinais manuais, entre outros recursos que serão discutidos a fundo nas seções sobre as práticas de listagem e contraste.

7.3.4. Piscada de olhos

Além do padrão de aceleração-desaceleração e dos gestos atencionais coesos – os dois fatores mais recorrentes nas instâncias analisadas –, um sinal não-manual que está claramente correlacionado às fronteiras de UEs é a piscada dos olhos. Tal fenômeno tem sido constatado no estudo de várias LSs (Baker e Padden, 1978; McCleary e Viotti, 2007; Nespor e Sandler, 1999; Wilbur, 1994, entre outros) e, como sugerem Nespor e Sandler (1999), a explicação parece ser a de que, assim como se dá com as inspirações de ar nas LOs, as piscadas nas LSs devam ser produto de uma necessidade fisiológica humana e que, ao co-ocorrerem com a fala, tendem a se submeter à sua organização gramatical, incrementando a própria estrutura da língua.

Uma das passagens que ilustra esse papel das piscadas aparece na figura 39,¹⁷² que traz a transcrição referente a um reparo iniciado por R. A frase, PROFESSOR EXPLICAR VOCÊ (ver discussão sobre alongamento final acima), envolve um gesto atencional típico de perguntas sim/não que revela uma forte coesão ao longo de todo o conteúdo segmental com o qual ele co-ocorre. Essa coesão é quebrada precisamente ao final do golpe do sinal VOCÊ, quando R dá uma forte piscada (“bb”, trilha R.4) acompanhada de um aceno de cabeça (que não aparece na transcrição). A confirmação de que tais marcas delimitam o término da unidade, então, se revela no fato de elas serem imediatamente seguidas de uma retração total do sinal VOCÊ (“rt”, trilha R.3).

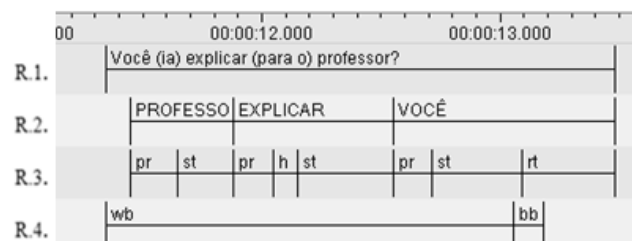


Figura 39. Piscada de olhos delimitando o final de uma unidade entoacional.

7.3.5. Acento enfático

Por fim, as instâncias de turnos simples me permitiram levantar ainda uma hipótese sobre a manifestação formal do acento enfático na libras. Numa das frases analisadas, chama a atenção as sobrancelhas erguidas e a forte projeção da cabeça para frente e para trás na realização do sinal IGUAL na frase VOCÊ OBRIGAR ESS@-DOIS IGUAL (ver discussão sobre alongamento final acima). No que se refere às modulações do sinal manual, nota-se uma longa preparação, uma manutenção da suspensão pré-golpe por quatro frames, seguida de um rápido deslocamento do golpe no espaço.¹⁷³

Se os graus de modulações da fala e da expressividade corporal que a acompanha estão iconicamente relacionados ao grau de importância que o falante confere às informações no discurso, como propõe Bolinger (1983: 160), a instância em questão pode ser vista como a marcação de uma informação com alto grau de

¹⁷² A trilha R.4, referente à pálpebra, envolve as seguintes siglas de notação: “wb” (*both eyes wide open*) e “bb” (*both eyes blink*).

¹⁷³ Sobre o deslocamento do sinal no espaço, cabe ainda assinalar que a interação do sinal manual com os não-manuais (i.e. a projeção da cabeça à frente durante a preparação de IGUAL, e para trás durante o golpe) dá a impressão de que o sinal percorreu um maior deslocamento no espaço – impressão que pode ser futuramente investigada com experimentos de percepção da fala junto a falantes da libras.

importância no discurso. De fato, tal proposta se justifica quando notamos que, no contexto em questão, R se mostra indignada com a persistência de W em exigir que seu colega se comporte de maneira *igual* a ele no curso de física, o que motiva o tom de seu questionamento.

Com essa hipótese sobre o acento, então, encerra-se a primeira parte da análise sobre as características formais de UEs na libras. A análise de turnos simples, possivelmente formados por uma única unidade, mostrou que aspectos como o padrão de aceleração-desaceleração (realizado por meio de modulações nos sinais manuais), os gestos atencionais coesos (envolvendo configurações do corpo que acompanham a sinalização), e as piscadas de olhos se destacam como pistas importantes na tarefa de segmentação do discurso. O próximo passo será o de olhar a estruturação de algumas práticas sociais convencionais, com o intuito não apenas de corroborar a aplicabilidade desses primeiros achados, mas também de identificar novos recursos relevantes para a questão.

7.4. Instâncias de turnos complexos estruturados

7.4.1. Listas

Na investigação do trecho do corpus transcrito, uma lista produzida por W foi inicialmente identificada e analisada em seus aspectos funcionais e formais, tomando como base de comparação a estruturação da prática de listagem em LOs. É essa lista que será objeto de discussão central da presente seção. Apesar disso, outras treze listas posteriormente identificadas no corpus – tanto na conversa entre R e W quanto nas outras duas conversas registradas – permitiram alcançar uma maior generalização da análise, de modo que outras listas e/ou aspectos relativos a elas deverão também ser trazidos na discussão.

Uma primeira dimensão a ser analisada refere-se à função das listagens, bem como à sua estruturação seqüencial na conversação em libras. A passagem abaixo traz uma *tradução* da lista, envolvendo o contexto mais amplo no qual ela se insere.¹⁷⁴

¹⁷⁴ Tendo em vista a dificuldade de se escrever ou transcrever o discurso em libras, optei, diante da necessidade de reproduzir excertos de discurso maiores, em trazer aqui uma tradução do trecho relevante a fim de facilitar a discussão.

Algumas observações sobre a tradução apresentada: i) a fala de ‘R’ aparece deslocada à direita para facilitar a sua visualização; ii) as aspas, em unidades como 2 (“Um trabalho?”), referem-se ao discurso

- 1 (W) Me falou que hoje eu ia vir explicar para a classe.
- 2 (W) "Um trabalho?"
- 3 (W) "Qual o tema?"
- 4 (W) "Física."
- 5 (R) Explicar o que?
(segue uma breve troca em que W soletra à sua interlocutora a palavra FÍSICA, cujo sinal era desconhecido por ela)
- 6 (W) Eu ia explicar. (W coça a cabeça em sinal de dificuldade)
- 7 (W) A leitura era difícil,
- 8 (W) Era complicado.
- 9 (R) Você ia explicar para o professor?
- 10 (W) É!
- 11 (W) Eu ia ser avaliad=
- 12 (W) Primeiro eu,
- 13 (W) Segundo o IL-2,¹⁷⁵
- 14 (W) O cara chamado Luis,
- 15 (W) Tinha esses dois e o terceiro era uma mulher ouvinte,
- 16 (W) Quarto era uma mulher ouvinte,
- 17 (W) Era o grupo.
- 18 (W) Eu vinha explicar o tema "física".
- 19 (W) Eu... (W coça a cabeça em sinal de dificuldade)
- 20 (W) Eu não conhecia aquelas palavras de significado difícil.

A <[lista produzida por W](#)> aparece assinalada pelas setas à esquerda. Observando o contexto em que ela se insere, vemos que a natureza *encaixada* da prática, proposta por Selting (2003), também caracteriza a listagem na conversa em libras. O contexto trata de uma avaliação de física que envolvia a apresentação de W perante a turma e as suas dificuldades com essa apresentação. A realização da listagem, com fins de reparo da frase "Eu ia ser avaliad=", é precedida e sucedida por esse contexto, o que se evidencia pela re-contextualização da lista nas linhas 18 a 20. Nela, W retoma com praticamente as mesmas palavras e gestos usados anteriormente (linhas 1 a 4 e linhas 6 a 8) a tarefa que fora incumbido de fazer e a sua dificuldade em realizá-la.

direto na fala de W; iii) os parênteses trazem descrições sobre aspectos do contexto considerados relevantes; iv) o sinal '=' indica interrupção abrupta da palavra.

¹⁷⁵ No mundo surdo, as pessoas recebem um nome na libras, um sinal, em geral associado a alguma característica física da pessoa. Esse nome acompanha a vida social da pessoa em interações sinalizadas, ao lado do nome original em português. No caso em questão, "I-L2" é o sinal do colega de W que, em seguida, teve o seu nome em português, "Luis", revelado.

Para Sanchez-Ayala (2003), as listagens desempenham basicamente duas funções interacionais dentro dos contextos mais amplos nos quais elas se inserem: estabelecer uma base de conhecimento comum (*common ground*) ou trazer suporte para um dado argumento. Das quatorze listas identificadas nas conversas em libras, seis se mostraram relacionadas à primeira função – entre elas a lista aqui discutida, que traz informações relevantes sobre a situação de avaliação de W – e as outras sete, à segunda função, servindo de suporte a argumentos correntes no discurso.¹⁷⁶

No que diz respeito à estrutura tripartida da lista em si, proposta por Jefferson (1990), a prática na libras parece mais uma vez corresponder aos achados referentes às LOs. No início da listagem, a ocorrência de uma bóia de listagem,¹⁷⁷ traduzida como “primeiro” na linha 12, permite uma projeção inicial da prática de detalhamento, ainda que não seja possível estimar quantos itens serão listados; esse detalhamento continua nas linhas 12 a 16, que envolvem itens pertencentes à lista e uma breve digressão na linha 14; por fim, ao término da listagem, na linha 17, “Era o grupo” traz uma expressão genérica que categoriza todos os itens listados, encerrando a prática. A “obediência” dessa lista à estrutura tripartida interna e à natureza encaixada das listas em outras línguas indica que a orientação a essa forma de estruturação da prática seja universal, tendo em vista que línguas como o inglês e o alemão, estudados por Jefferson (1990) e Selting (2003), e a libras, aqui analisada, não possuem qualquer relação histórica.

Passando à análise dos recursos formais utilizados na construção das listagens, um primeiro passo a ser tomado é o de verificar de que maneira um falante surdo demonstra ao outro que uma prática desse tipo foi iniciada, está em curso e foi finalizada, bem como qual é o estatuto dos elementos que aparecem no corpo da lista

¹⁷⁶ Cabe assinalar que essa distinção não me parece tão clara quanto Sanchez-Ayala sugere. Numa das conversas, em particular, um dos participantes fala ao seu interlocutor, um instrutor de libras novato, sobre os vários anos de experiência com o ensino dessa língua. Nesse contexto, as listas produzidas sobre aspectos dessa experiência podem ser interpretadas, a meu ver, de duas maneiras: mais objetivamente, como informações contextuais sobre essa experiência; ou, retoricamente, como argumentos revelando o quão experiente ele é.

Além disso, uma das listas do corpus, relacionada ao estabelecimento de uma base de conhecimento comum, foge levemente do padrão encontrado nas demais. Trata-se de uma ocorrência em que o falante S representa, por meio do discurso direto, a própria ação de leitura de uma listagem (os itens de seu currículo) por um empregador, numa entrevista de emprego reportada a seu interlocutor (ver discussão na nota de rodapé 184).

¹⁷⁷ Bóias nas LSs (Liddell, 2003) são sinais produzidos com a mão passiva que são mantidos parados no ar, numa dada configuração, enquanto a mão ativa continua a produzir outros sinais. As bóias *de listagem* constituem apenas um dos vários tipos de bóias existentes nessas línguas.

(i.e. como itens listados, ou como digressões). Tendo em vista que as diferentes partes da lista e a própria lista como um todo constituem unidades lingüísticas de diferentes níveis, ao se levantar os recursos prosódicos que permitem essa identificação, estaremos conseqüentemente levantando aspectos formais para a segmentação do discurso sinalizado em unidades gramaticais.

Nas LOs, Selting (2003) mostra que a estruturação das listagens é feita principalmente pelo uso de padrões entoacionais, ainda que recursos sintáticos e lexicais também sejam importantes. A prosódia parece ser o *principal* recurso porque, como mostra a autora, quando recursos lexicais e sintáticos competem com a prosódia sobrepondo-se na construção das listas, é a prosódia que “vence” a disputa e que orienta, em última instância, o modo como os participantes interpretam o papel de um dado evento de uso na organização da seqüência de fala imediata. Essa orientação é verificada por meio da análise de como e quando os participantes respondem às partes componentes da lista e à prática como um todo.

A figura 40 abaixo traz uma transcrição simplificada da lista, acompanhada de fotos dos sinais.¹⁷⁸ Do ponto de vista lexical, um primeiro recurso importante de estruturação da prática que pode ser identificado é a bóia de listagem. Como revela a pauta 2, a realização do item lexical BÓIA-UM com a mão esquerda (‘me’), para onde W aponta com a mão direita por meio do sinal ESS@, é o recurso que permite uma projeção de que tal prática está em curso.¹⁷⁹ Desde que emerge, na pauta 2, a bóia permanece suspensa no espaço de sinalização enquanto a mão direita de W realiza outros sinais, até que o último item listado, na pauta 6, seja realizado. O término da lista, então, é indicado por um outro item lexical, o finalizador GRUPO, que emerge, na pauta 7, como uma generalização de todos os itens previamente listados.

¹⁷⁸ Essa combinação de transcrição e fotos foi adotada a fim de facilitar a apreensão das informações relevantes para a análise. Cada uma das pautas refere-se a unidades prosódicas postuladas na análise. As convenções adotadas na notação serão apropriadamente mencionadas quando a sua discussão se mostrar pertinente.

¹⁷⁹ Embora o sinal BÓIA-UM seja formalmente idêntico ao numeral UM, há evidências, relativas a diferenças de forma e função, de que as bóias de listagem devam ser entendidas como um recurso distinto dos numerais – um possível caso de homonímia na libras.








 <p>1. EU AVALIAR= Eu (ia ser) avaliad=</p>	 <p>2. ESS@ MIM me BOIA-UM..... Primeiro eu,</p>
 <p>3. ESS@ I-L2:: me BOIA-DOIS..... Segundo o I-L2,</p>	 <p>4. >HOMEM >NOME L-U-I-S me BOIA-DOIS..... (O) cara chamado Luis.</p>
 <p>5. ESS@ ESS@ +ESS@ >MULHER >OUVINTE me BOIA-DOIS.....BOIA-TRÊS..... (Tinha) esses dois (e o) terceiro (era uma) mulher ouvinte,</p>	
 <p>6. ESS@ >MULHER >OUVINTE me BOIA-QUATRO..... Quarto (uma) mulher ouvinte.</p>	 <p>7. ^.. GRUPO:: (Era o) grupo.</p>

Figura 40. Instância de listagem na libras.

A consideração das demais instâncias de listagem permitiu a identificação de outros recursos lexicais relevantes para a construção dessa prática na libras – análise que deve ser aprofundada futuramente com uma ampliação do corpus. A bóia de listagem, por exemplo, quando inexistente, pode ceder lugar a outros itens lexicais que, por sua recorrência a cada novo item listado, também contribuem para a identificação do corpo da lista e para a coesão da prática como um todo. Entre esses recursos estão

conectivos como <TAMBÉM> ou <MAIS> e verbos que entram na construção dos itens listados (e.g. o verbo <DAR>, que, numa lista sobre os benefícios de um emprego, recorre em frases do tipo ‘te dão isso, te dão aquilo’; ver discussão relativa à figura 44).

Em relação aos finalizadores, como o sinal GRUPO, nota-se que, embora não pareçam ser obrigatórios, eles são de fato recorrentes ao término das listagens. Como Selting (2003) argumenta, tais recursos envolvem itens lexicais genéricos que, ao término da lista, retomam os elementos previamente listados, categorizando-os e, desse modo, contribuindo para uma maior coerência da prática. Por esse motivo, embora freqüentemente envolvam palavras genéricas, como <VÁRIAS-COISAS>, os finalizadores se revelam em muitos casos dependentes do conteúdo semântico dos itens enumerados.¹⁸⁰

No que se refere aos sinais não-manuais, um dos gestos atencionais cruciais na estruturação da prática da listagem na libras é o olhar. Dois tipos de olhar puderam ser identificados na lista de W da figura 40: o olhar para a bóia de listagem e o olhar para o interlocutor. Como mostra a tabela 3 abaixo – na qual o olhar para a bóia aparece representado pelo fundo cinza e o olhar para o interlocutor, pela ausência de cor ao fundo – W intercala esses dois tipos de olhares durante a sinalização de maneira bastante consistente, produzindo um claro padrão paralelístico.

Item Enumerado	Detalhamento
ESS@(1)	MIM
ESS@(2)	JL-2 HOMEM NOME J-U-A-R-E-Z
ESS@(2) ESS@(1) ESS@(3)	MULHER OUVINTE
ESS@(4)	MULHER OUVINTE

Tabela 3. Paralelismo no direcionamento do olhar durante a prática da listagem.

Esse padrão do olhar acompanha o paralelismo sintático das enumerações e de suas respectivas descrições. Quando W sinaliza as bóias, seu olhar se desloca para elas; quando sinaliza o conteúdo das enumerações, retoma o contato visual com R. Duas instâncias fogem a esse padrão, que pode ser tomado como análogo à exploração paralelística de tons terminais nas listagens de LOs: em primeiro lugar, durante a sinalização de MIM, o olhar de W não retorna à R, permanecendo voltado para as mãos;

¹⁸⁰ Por exemplo, na lista em que S reporta a leitura de um currículo, os vários pontos são retomados ao final por meio de inúmeras iterações do sinal <ITEM>, e numa lista sobre os grupos que compõem o público de uma palestra, o conjunto do público é referido ao final por <TODOS PESSOAS-EM-LINHAS>.

em segundo lugar, em meio à soletração de “Luis”, o olhar se desvia brevemente para as mãos e, em seguida, retorna à R.

No primeiro caso, relativo ao sinal MIM, é difícil dizer com segurança – no atual estágio de investigação da libras – por que razão W manteve o olhar para as mãos (pauta 2, figura 40). Todas as outras instâncias de listas do corpus que contêm bóias revelam o direcionamento do olhar para o interlocutor quando o conteúdo das enumerações é descrito. Uma possível razão dessa ausência de contato visual é o fato de o sinal MIM, em questão, fazer referência a um elemento pressuposto (W já havia dito “Eu ia ser avali-” antes de iniciar a listagem), isto é, um elemento que podia ser mencionado sem um monitoramento da reação de R por parte de W. Nesse aspecto, porém, a análise de um corpus mais abrangente se torna necessária para a postulação de uma hipótese menos especulativa.

A segunda exceção ao padrão gestual paralelístico de W na construção de sua listagem é o breve desvio do olhar durante a soletração de “Luis” (pauta 4, figura 40). Considerando-se que esse desvio é bastante comum de ser observado nas soletrações manuais, tal fenômeno poderia ser considerado como parte convencional dessa prática nas LSs. Contudo, embora a grande maioria de soletrações manuais envolva esse tipo de desvio de olhar, algumas são realizadas sem eles. Esse fato, aliado ao pressuposto teórico desta tese sobre a função atencional dos gestos, levou-me à hipótese de que o breve deslocamento do olhar, do interlocutor para as mãos, tenha como função direcionar a atenção do interlocutor para uma expressão particularmente relevante, conferindo-lhe uma maior saliência. Tal recurso pode ser visto como funcionalmente análogo ao acento nas LOs.¹⁸¹

A hipótese é a de que o efeito de salientação desse desvio do olhar não deva ser resultado da sua simples mudança de um local para outro, tendo em vista que o olhar é

¹⁸¹ Ainda assim, permanece a questão de por que as soletrações manuais, de maneira geral, envolvem breves desvios do olhar para as mãos. Uma possibilidade é a de que a soletração de palavras esteja rotineiramente associada a informações de maior relevância interacional, conseqüentemente exigindo recursos de salientação de maneira igualmente habitual. Outra possibilidade é a de que a soletração manual em si, por ser uma prática que exige um grande refinamento do olhar tendo em vista a sutileza de seus movimentos distintivos, favoreça o emprego de recursos de salientação a fim de facilitar a sua apreensão perceptual pelos falantes de libras – exceto nos casos em que a soletração se torna óbvia pelo contexto ou pela simplicidade da palavra soletrada. Seja qual for a motivação, de salientação informacional e/ou perceptual, a hipótese sobre a função atencional dos breves desvios do olhar durante a sinalização pode ser mantida.

um dos articuladores mais dinâmicos na libras. Mais especificamente, o que parece causar esse efeito é a brevidade com que o olhar é deslocado *dentro de uma dada unidade gramatical*, divergindo de um padrão estável em que um ponto X é focado para um outro ponto, Y, no curso da sinalização, e retornando rapidamente ao ponto X. Esse deslocamento, quando dentro de uma unidade, rompe com a estabilidade atencional esperada e acaba por salientiar o conteúdo segmental sob escopo do olhar divergente.

Além do olhar, que por seu padrão paralelístico permite não apenas a segmentação de unidades mas também a salientação de elementos dentro delas, outro recurso de segmentação identificado nessa lista é a expressão facial. Tal recurso, que aparece na produção da frase “HOMEM NOME LUIS” (pauta 4, figura 40), é caracterizado neste contexto pela combinação de olhos semi-serrados e bochechas contraídas, formando um leve sorriso. A expressão perdura pelo tempo exato que W leva para produzir a frase em questão, como mostra a figura 41 abaixo.

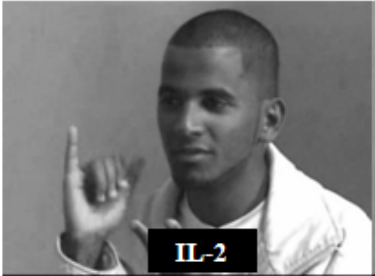


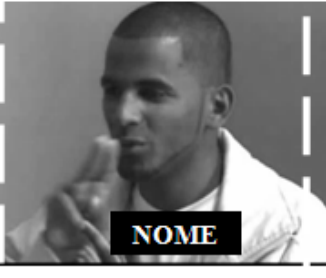

DESCRIÇÃO DO ITEM 2	DIGRESSÃO	RETOMADA DA LISTA
 <p>IL-2</p>	 <p>HOMEM</p>	 <p>ESS@</p>
<p>A expressão facial OLHOS SEMI-CERRADOS E BOCHECHAS CONTRAÍDAS, que aparece nas três figuras centrais destacadas ao lado, tem uma dupla função semântica: <i>pragmaticamente</i>, evoca algo supostamente familiar ao interlocutor; <i>prosodicamente</i>, atua no nível informacional, enquadrando o conteúdo segmental sobre o qual tem escopo e contribuindo, portanto, para o direcionamento da atenção do interlocutor.</p>	 <p>NOME</p>  <p>LUIS</p>	<p>Comparando esse recurso da libras com os recursos prosódicos das línguas orais, tais expressões podem ser entendidas como similares, por exemplo, às mudanças no nível geral de volume e pitch. Fonologicamente, elas envolvem alterações de certos parâmetros prosódicos de uma UP para a outra; semanticamente, trazem, além de seu sentido pragmático, a função de agrupar um dado conteúdo segmental para ser simultaneamente visualizado num único quadro atencional.</p>

Figura 41. Expressão facial como um tipo de gesto atencional na libras.

O conteúdo segmental que essa expressão facial agrupa não é interpretado como um elemento novo, listado por W; trata-se de uma elaboração do item anterior, “I-L2”, talvez porque W imaginou que a sua interlocutora poderia não conhecer ou se lembrar da pessoa referida por esse sinal. De fato, pelo meu conhecimento da libras, esse tipo de expressão facial é utilizado justamente quando se pretende evocar algo familiar no interlocutor (i.e. algo como: “Você sabe/conhece isso que eu estou dizendo?”).¹⁸²

Ao agrupar esse conteúdo específico, a expressão facial funciona, entre outras coisas, como um gesto atencional coeso que destaca a referida digressão (i.e. a elaboração sobre um dos elementos listados) da prática de listagem em curso. Desse modo, além de sua função pragmática, pode-se dizer que, semanticamente, esse gesto sirva para delimitar o conteúdo segmental com o qual ele co-ocorre, situando-o num quadro atencional separado.

Além do olhar e da expressão facial, um outro recurso que se mostrou relevante para o gerenciamento da atenção na construção da listagem em libras foi a modulação de sinais manuais. Um primeiro tipo de modulação, representado na transcrição da figura 40 por meio da anotação ‘>’, é a redução fonética dos sinais. Tal como discutido na seção 7.3.2, esse fenômeno se manifesta aqui em diferentes contextos frasais, funcionando como indicador de pontos de *disjunção mínima* na cadeia de fala. Na lista produzida por W, esses pontos aparecem nos sinais HOMEM e NOME (pauta 4) e MULHER e OUVINTE (pautas 5 e 6), que têm o número de repetições internas ao golpe reduzido, passando de dois, relativos à forma de menção, para um, nos sinais em contexto.

Outro tipo de modulação importante – também identificada anteriormente, na seção 7.3.1 – é o alongamento final, representado na transcrição da figura 40 por meio da anotação ‘::’. Tal alongamento envolve, no caso do sinal I-L2 (pauta 3), a manutenção da suspensão pós-golpe por oito frames, e no caso do sinal GRUPO (pauta 7), o aumento das repetições internas ao golpe, que passam de duas, relativas à sua forma de menção, para três, no sinal em contexto. Esses dois alongamentos finais se

¹⁸² Tal expressão é idêntica à que Sandler e Lillo-Martin (2001) reportam como a expressão de “informação assumida como compartilhada” na língua de sinais israelense, ISL (p. 10). Essa curiosa similaridade entre recursos gramaticais de línguas tão distintas quanto a libras e a ISL levanta a questão de em que medida a gramática das LSs não poderia ser mais bem compreendida considerando-se a base gestual universal que lhe serve de sustentação.

mostram especialmente significativos quando nos damos conta de que os locais onde eles ocorrem constituem fronteiras importantes na seqüência interacional: a primeira, envolvendo a transição entre a prática de listagem em si e uma breve digressão; e a segunda, envolvendo a transição entre o término da listagem e a sua recontextualização na narrativa dentro da qual ela emergiu.

Essa correlação entre rupturas maiores no fluxo de informação e rupturas maiores na fluidez da fala vem corroborar um dos achados de Chafe (1980: 40-47). Em sua análise das “Histórias da Pêra”, o autor mostra que quebras no fluxo da narrativa (e.g. a inserção de eventos não diretamente relacionados ou relevantes para o desenvolvimento do enredo central) se refletem diretamente na fala dos participantes por meio de rupturas formais mais salientes. Tais marcas seriam reflexo da necessidade de um maior investimento cognitivo por parte do falante em pontos onde uma reorientação mais ampla no *centro de interesse* se faz necessária.¹⁸³

O fato de que rupturas maiores no fluxo de informação trazem reflexos na fala de W é corroborado também pela unidade que aparece na pauta 5 da figura 40 (“Tem esses dois e o terceiro era uma mulher ouvinte”). A breve digressão de W para elaborar o referente “I-L2”, por meio da frase “O cara chamado Luis”, exige, por menor que possa parecer, que W recontextualize a lista então interrompida antes de dar continuidade a ela. Ele faz isso por meio da retomada dos dois itens anteriormente citados, apotando para o segundo e o primeiro itens de BÓIA-DOIS, antes de passar ao item BÓIA-TRÊS. Interessantemente, essa retomada da lista – que, por si só, constitui-se numa outra prática de listagem – apresenta um terceiro tipo de modulação de sinais manuais relevante para a presente discussão.

Essa “sublista” produzida por W envolve a enumeração de três itens: os dois previamente listados e um terceiro, a ser tratado a partir dali. No que diz respeito aos sinais manuais, os dois sinais referentes às informações dadas (i.e. ESS@ e ESS@, direcionados respectivamente ao segundo e ao primeiro item de BÓIA-DOIS), são

¹⁸³ O fato de as listas ocorrerem freqüentemente como práticas encaixadas em contextos mais amplos faz com que, especialmente quando a listagem é longa, uma recontextualização da narrativa após a listagem se faça necessária – como na lista de W acima discutida. Nesse tipo de situação, em mais de uma instância de listagens identificadas, o item finalizador da lista é não apenas acompanhado de um alongamento final maior, como apontado acima, mas também por um aceno marcado da cabeça ao término da listagem. Na lista de W, esse aceno pode ser observado na pauta 7 da figura 40, pela notação ‘^’ na trilha da cabeça, imediatamente acima da trilha das glosas.

realizados rapidamente, com uma fase de preparação e uma fase expressiva bastante breves. Já o terceiro sinal, referente ao elemento novo da lista que deverá ser então elaborado (i.e. ESS@, direcionado ao terceiro item de BÓIA-TRÊS), é realizado com uma fase de preparação significativamente mais longa, uma suspensão pré-golpe de dois frames e uma fase expressiva que percorre uma distância maior no espaço quando comparada à fase dos dois sinais anteriores. O fato de os três sinais manuais serem fonologicamente idênticos (ESS@ é um gesto de apontamento, nesse contexto direcionado à bóia de listagem) sugere fortemente que a modulação no terceiro sinal, associada à informação nova no seu contexto imediato, seja uma espécie de *acento* da palavra que lhe confere proeminência.

No que diz respeito aos sinais não-manuais, a análise dessa pequena lista revela ainda um padrão paralelístico distinto do padrão do olhar identificado na listagem maior, na qual ela se encaixa. Simultaneamente à realização do pico de cada um dos sinais ESS@, a cabeça e o tronco de W se inclinam sutilmente para a frente, retornando à posição neutra ao término do sinal – movimento representado pela anotação “v” colocada acima de cada instância do sinal ESS@ (pauta 5, figura 40) (ver também Liddell, 1980). Essas idas e vindas acompanham inclusive a modulação dos sinais manuais, com a cabeça inclinando-se para trás de uma maneira mais intensa durante a preparação também mais intensa do terceiro sinal ESS@.

Esse paralelismo gestual, mais tarde, pôde ser observado na construção de outras instâncias de listagem no corpus, constituindo-se, portanto, num recurso alternativo ao paralelismo do olhar. Por exemplo, <[na passagem abaixo](#)>, R está respondendo à fala de seu interlocutor, S, que havia manifestado seu desejo de conseguir uma barriga “tanquinho” e de ficar bem fisicamente. A lista produzida por R mostra a sua reprovação em relação a esse tipo de preocupação, ressaltando as coisas que ela acha realmente importante de serem buscadas.

<div style="text-align: center;"> >... </div> <div style="text-align: center;">1. AH PESSOA_x</div> <div style="text-align: center;">Ah, (uma) pessoa</div>		<div style="text-align: center;">>... .>.....</div> <div style="text-align: center;">2. SE OLHAR_{x,y} PESSOA_y</div> <div style="text-align: center;">Se (ela) olha (pra) mim</div>	
<div style="text-align: center;">v</div> <div style="text-align: center;"></div> <div style="text-align: center;">3. IMPORTANTE AMOR</div> <div style="text-align: center;">(O) importante (é o) amor,</div>	<div style="text-align: center;">v</div> <div style="text-align: center;"></div> <div style="text-align: center;">4. RELAÇÃO IMPORTANTE</div> <div style="text-align: center;">(A) relação (é) importante.</div>	<div style="text-align: center;">v</div> <div style="text-align: center;"></div> <div style="text-align: center;">5. SENTIMENTO</div> <div style="text-align: center;">(O) sentimento,</div>	

Figura 42. Paralelismo nos movimentos da cabeça e do tronco na prática de listagem.

A trilha imediatamente superior à da glosa refere-se ao olhar. Como pode ser visto pela anotação ‘|’, que representa o olhar direcionado ao interlocutor, a partir do momento em que a lista se inicia (pauta 3), R mantém o seu contato visual com S inalterado até o final (pauta 5), diferentemente do padrão paralelístico do olhar nas listas que utilizam bóias. Por outro lado, como mostra a anotação “V” na trilha da cabeça, imediatamente acima da trilha do olhar, esse recurso é substituído por um paralelismo no movimento da cabeça, que se inclina para frente durante a realização de cada um dos itens listados (i.e. AMOR, RELAÇÃO e SENTIMENTO). Nessa variante, portanto, o padrão paralelístico característico das listagens é realizado por meio de um recurso distinto daquele encontrado em listas com bóias.

Uma observação interessante nessa lista de R é a de que o último item listado, SENTIMENTO, acaba sendo realizado somente após o término da lista. Esse término se revela pela produção do item finalizador, IMPORTANTE, que, nesse caso, serve tanto como disparador da prática de detalhamento, ao início da lista (pauta 3), quanto como elemento genérico que categoriza os itens listados ao término da mesma (pauta 4). Além disso, a identificação de IMPORTANTE e SENTIMENTO como, respectivamente, finalizador e novo item listado se revela também pela ausência de inclinação da cabeça para a frente na realização do primeiro sinal, e pela retomada dessa inclinação (i.e. a retomada do padrão paralelístico dos itens listados) na realização do segundo sinal.

Tal análise é corroborada pela fala subsequente de S, o interlocutor de R, em resposta à lista. Como mostra a figura 43 abaixo, tão logo R termina a sinalização de IMPORTANTE (trilha R.2, 9m12s600ms), S dá início à sua resposta. Isso evidencia que S de fato identificou o sinal IMPORTANTE como o finalizador da lista, em especial considerando a ausência de inclinação da cabeça à frente na realização desse sinal. Contudo, como já afirmado, R opta, nesse preciso momento, por acrescentar um novo item à lista, o que acarreta uma sobreposição entre esse novo item, SENTIMENTO, e o início do turno de S.

É interessante, nesse sentido, notar a importância do padrão paralelístico na construção das listagens. Ocorre que a retomada da inclinação da cabeça à frente em SENTIMENTO permite que, mesmo estando *fora da listagem*, esse elemento seja devidamente integrado por S como *parte da listagem*. Essa análise é mais uma vez corroborada pela fala subsequente de S que aparece [<na passagem a seguir>](#), na qual ele

retoma a lista de R – a fim de contrastá-la com um outro argumento –, incluindo o item SENTIMENTO entre os elementos listados (trilha S.2, 9m13s900ms).

	09.000	00:09:10.000	00:09:11.000	00:09:12.000	00:09:13.000	00:09:14.000	00:09:15.000	00:09:16.000					
R.1	Ah (uma) pesso Se (ela) olha (para) (O) importante (é o) am (A) relação (é) important (O) sentimento.												
R.2	AH	PESSOA	SE-EL	OL	PESSO	IMPORT	AMOR	RELAÇÃO	IMPORT	SENTIMENT	AH		
S.1						Ah, isso!	Espera, eu sei!	Sentimento, amor, é importante.	A relação (tá) cert				
S.2						ISSO	FIC EU	SABE	SENTIMENT	AMOR	IMPO	RELAÇÃ	CERTO

Figura 43. Importância do paralelismo como elemento de coesão das listas.

Essa análise também corrobora a afirmação de Selting (2003) de que, quando recursos prosódicos e léxico-sintáticos competem na construção da listagem para sinalização de sua organização interna e de suas fronteiras, é a prosódia que vence a disputa. No caso em questão, o item SENTIMENTO aparece como seqüencialmente (i.e. sintaticamente) posterior ao elemento finalizador da listagem, mas a retomada do padrão paralelístico, por meio da inclinação da cabeça à frente na realização desse item, permite que ele seja devidamente integrado ao corpo da lista, como um item deslocado de lugar.¹⁸⁴

¹⁸⁴ Esses dois padrões gestuais paralelísticos de organização das listagens encontram ainda duas pequenas variações. No que se refere ao padrão dual do olhar, além do “olhar para *bóia* ↔ olhar para o *interlocutor*”, o corpus traz também instâncias de “olhar para o *lado* ↔ olhar para o *interlocutor*”. No que se refere ao padrão dual da cabeça-tronco, além da “inclinação para *frente* ↔ inclinação para *trás*”, o corpus traz também uma instância de “inclinação para a *direita* ↔ inclinação para a *esquerda*”. Por fim, há ocorrências de listas com bóias que envolvem simultaneamente um padrão dual de cabeça-ombro e de olhar. A seleção entre um ou outro padrão não deve ser arbitrária, mas motivada, de modo que razões para a motivação de o falante adotar uma outra outra estruturação devem ser buscadas.

Nesse aspecto, a combinação da análise de dados espontâneos com a de experimentos de eliciação, tal como aqueles adotados em estudos laboratoriais da prosódia, me parece especialmente frutífera. Pode-se pedir aos informantes, por exemplo, que respondam a perguntas que eliciem listas abertas (e.g. o que você faz nos fins de semana?), listas fechadas (e.g. quantos irmãos você tem?), entre outros tipos possíveis, e verificar quais são os padrões empregados.

Embora não seja meu intuito aqui analisar especificamente essas diferenças, tendo em vista o foco atual sobre a questão da segmentação do discurso, gostaria de tecer uma breve consideração a respeito da iconicidade dos gestos atencionais e da sua possível relação com a escolha de determinados recursos gramaticais na construção das listagens em libras.

Em <[uma das listagens produzidas](#)>, S, por meio do discurso direto, assume a voz de um empregador numa entrevista de emprego da qual ele próprio participou como candidato – e que ele agora reporta a W. No espaço enunciativo da narrativa, então, o personagem-empregador segura o *currículo vitae* do personagem-candidato em sua mão esquerda, passando por cada um de seus itens e lendo-os para o candidato à sua frente. Esse procedimento de leitura se desenvolve de maneira bastante similar à lista produzida por W, analisada na figura 40: S alterna o olhar sistematicamente, primeiro para a mão esquerda, onde o currículo é “segurado”, e em seguida para o interlocutor “à sua frente”.

O paralelo formal entre a ação de “ler uma lista em voz alta para alguém” e a ação de “produzir uma

Por fim, um último recurso importante para a segmentação do discurso identificado na análise das listas é a piscada do olhar. Em todas as listas, as piscadas aparecem com relativa frequência ao término de uma ou mais de suas unidades componentes. <[Uma das listas identificadas](#)>, porém, ilustra melhor do que quaisquer outras a relevância desse recurso.

Como mostra a figura 44 abaixo, ao término de três dos quatro itens listados por S (terceira imagem das pautas 1 a 3), a piscada de olhar aparece na fronteira final da unidade, revelando que, embora não obrigatório, tal recurso é recorrente e de grande relevância para a segmentação do discurso na libras.¹⁸⁵

Como a análise desta seção pôde revelar, portanto, a combinação de um olhar sobre as unidades mínimas – aproveitando-se do conhecimento sobre as características das UEs e dos gestos atencionais – com um olhar sobre práticas estruturadas – aproveitando-se do conhecimento sobre as listagens – pode ser uma forma bastante útil de levantar recursos de segmentação prosódica na libras. Na seção a seguir, a análise de uma outra prática recorrente, o contraste, permitirá o levantamento de ainda outros recursos.

lista numa interação em libras” salta aos olhos – a lista de S se situando num limiar entre a ação concreta e a ação reportada. Recuperando a lista de W, é como se ele, ao olhar para bóia, estivesse visualizando a lista propriamente dita e, ao olhar para sua interlocutora, R, reportasse o conteúdo de seus itens recém averiguados.

Sob uma perspectiva cognitivista, cabe destacar que, enquanto a lista que um empregador segura em mãos é um objeto físico, a lista produzida por W é um objeto cognitivo, para o qual não é possível olhar. A bóia de listagem e o direcionamento do olhar, nesse sentido, se revelam como âncoras materiais altamente icônicas para a produção ordenada da enumeração de itens na fala espontânea. Sob uma perspectiva linguístico-interacional, podemos sugerir a hipótese de que a iconicidade do padrão dual “olhar para a bóia \leftrightarrow olhar para o interlocutor” esteja correlacionada com a produção de listas fechadas, as quais os falantes dão início já com uma estimativa sobre os seus itens componentes e o seu possível ponto de completude.

¹⁸⁵ Essa listagem mostra ainda de que modo as inclinações de cabeça são coordenadas com os itens listados quando tais itens envolvem unidades compostas por mais do que um único sinal. Como mostra a trilha acima da dos olhos, referente à cabeça, inclinações com dois níveis de intensidade são produzidas por S: uma primeira, sobre o último sinal do que parece ser um sintagma (anotada com ‘v’, menor) e uma segunda, sobre o item lexical final da unidade (anotada com ‘V’, maior).

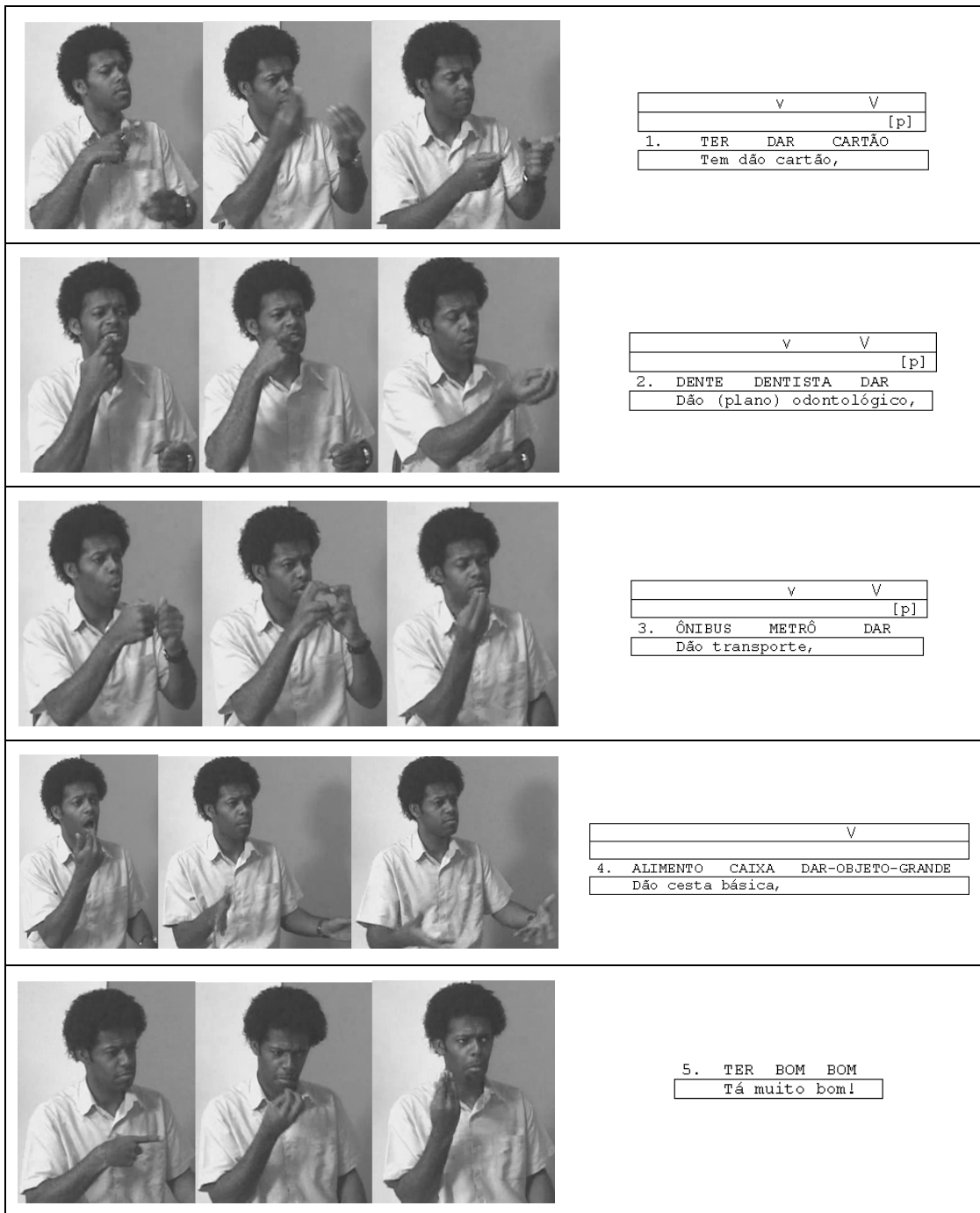


Figura 44. As piscadas do olhar delimitando unidades numa listagem.

7.4.2. Contrastes

Na busca de novas instâncias de listagem, deparei-me no corpus com alguns casos que, embora se assemelhem a essa prática em vários aspectos, envolviam apenas dois itens elaborados com fins de contraste, o que me levou a descartar a pertinência do termo “listagem” para essas ocorrências. A opção foi a de chamar tais práticas de *contrastes*, uma escolha que se revelou feliz tendo em vista que, sob essa denominação,

pude encontrar posteriormente outro trabalho da AC voltado para a análise do mesmo fenômeno (Atkinson, 1984). Essa leitura, por sua vez, me permitiu ter uma visão mais ampla das diferentes formas de estruturação da prática em questão, nem todas elas similares a listas, de modo que, ao voltar aos dados, pude identificar um total de quatorze instâncias de contrastes.

Os primeiros contrastes observados no corpus que me despertaram o interesse traziam uma grande semelhança estrutural com as listas. Sua organização interna parecia também envolver uma estrutura tripartida, formada por: a) a abertura do contraste, por meio de uma frase que incluísse algum item projetor; b) o corpo do contraste, envolvendo a elaboração de cada um dos dois itens contrastados; e c) um item finalizador, na forma de uma conclusão ou ainda de uma questão que transferia a responsabilidade da conclusão ao interlocutor. Além disso, os dois itens do contraste podiam ser opcionalmente projetados já no início da prática por meio do item lexical BÓIA-DOIS – o que mostra que o termo *bóia de listagem*, adotado por Liddell (2003), deva ser substituído por um termo mais genérico, como *bóia de enumeração*, tendo em vista a sua utilização em práticas outras que não as de listagem.

Com relação aos aspectos funcionais do contraste, tal prática parece de fato estar intimamente relacionada à busca de persuasão e convencimento do interlocutor, tal como propõe Atkinson (1984). A conversa do corpus na qual emerge o maior número de contrastes, entre R e W, é em grande parte marcada pelo caráter antagonístico da fala de W, que contrapõe a sua participação na escola à de um colega, Luis, sob uma perspectiva altamente crítica (i.e. W demonstra interesse, estuda, se dedica, enquanto Luis se omite, assume poucas responsabilidades, se mostra desinteressado, etc). Nesse contexto, os contrastes são utilizados por ambos os interlocutores: por W, na tentativa de marcar essa diferença; e, principalmente, por R, a fim de persuadir seu interlocutor de que ele estava errado no julgamento de seu colega.¹⁸⁶

¹⁸⁶ Embora essa prática se revele particularmente adequada em contextos polêmicos, como estratégia de persuasão – daí a sua alta relevância nos discursos políticos estudados por Atkinson (1984) –, uma variante dessa função básica encontrada no corpus é a exploração do contraste para provocação. Em uma das instâncias, S convida sua colega R para ajudá-lo numa palestra, e, ao explicar a ela a melhor maneira de chegar ao local, sugere que seja mais fácil fazer isso de trem. Diante disso, em tom jocoso, R responde com um contraste que, no português, poderia ser traduzido como: “Eu nunca [pego trem]. Te digo que sou rica. *Você é pobre, mas eu sou rica!*” (ver discussão nota de rodapé 190). Diante da expressão facial suspeita de S, então, R rapidamente acrescenta “Estou brincando!”. Essa justificativa mostra que o caráter polêmico dos contrastes estava latente também ali, mas, naquele contexto, não deveria ser levado a sério

Atkinson (1984) aponta que a prática do contraste se revela uma estratégia discursiva altamente eficaz na eliciação de respostas do público (i.e. aplausos) nos discursos políticos. Essa eficácia se revela também no corpus de conversas em libras, embora a resposta eliciada nesse gênero, obviamente, seja uma outra fala e não aplausos. No corpus, a grande maioria dos contrastes é imediatamente seguida de uma resposta que tem como intuito demonstrar concordância ou discordância em relação ao seu argumento implícito.

Essa observação nos conduz a uma reflexão sobre quais seriam os recursos estruturados dessa prática que permitem ao(s) interlocutor(es) identificá-la, projetando o seu término possível e, desse modo, podendo responder a ela de maneira temporalmente coordenada. A tabela 4 abaixo traz um dos contrastes mais complexos e bem estruturados do corpus. <Tal instância> se revela particularmente relevante como ponto de referência da análise por envolver vários recursos em combinação.

Glosas	Tradução
➡ 01 (R) ENTÃO DEPENDER PESSOA PESSOA	Ah depende de um e de outro.
➡ 02 (R) UM::	Uma pessoa diz,
➡ 03 (R) EU GOSTAR +PREFERIR+ ISSO::	“Eu gosto mais disso”.
➡ 04 (R) QUÍMICA	“Química”.
➡ 05 (R) >EU PROJETO::	“Eu faço um projeto”.
➡ 06 (R) AQUI	Esse é um.
➡ 07 (R) OUTR@ QUÍM= ISS=	Já outra qu= is=
➡ 08 (R) EU +PREFERIR+ ISSO QUÍMICA	“Eu não gosto disso química”.
➡ 09 (R) >DEPENDER >EU >FAZER +SIMPLES+ PALAVRA PALAVRA	“Talvez eu faça algo simples de dois conceitos”.
➡ 10 (R) DEPENDER PESSOA PESSOA PESSOA	Depende de cada pessoa.

Tabela 4. Instância de contraste complexo e altamente estruturado.

pelo receptor.

Essa passagem mostra também que a conversação, em especial, pode *estar* polarizada, mas não *é* polarizada. Isso significa que algum tipo de antagonismo pode emergir rapidamente numa dada seqüência conversacional e, logo em seguida, desaparecer, dependendo de como os participantes lidam com isso, como algo problemático ou não. Numa das conversas, por exemplo, S comunica W sobre uma palestra que ele irá fazer. Em meio às explicações sobre os arranjos necessários à palestra, S acrescenta um comentário que poderia ser traduzido como: “Eu vou ter que estar bem vestido. *Você indo de qualquer jeito tá bom, mas eu preciso estar chique!*”. É possível imaginar que uma colocação desse tipo desse início a uma discussão polarizada, possivelmente envolvendo o uso de novos contrastes; contudo, W, repetindo palavras do contraste de S e acenando positivamente com a cabeça, mostra que concorda com o seu interlocutor, reduzindo o antagonismo ao contexto específico onde ele emerge.

De modo similar às listas, o contraste em questão apresenta uma estrutura tripartida: i) uma unidade que contém uma afirmação genérica e que projeta a prática; ii) o corpo do contraste, formado por dois elementos elaborados; e iii) uma unidade que encerra a prática, também por meio de uma afirmação genérica. Essas partes são diferenciadas pelos diferentes tipos de setas na tabela 4: as brancas com fundo preto indicam a abertura e o fechamento do contraste (linhas 1 e 10); as pretas com fundo branco, o corpo do contraste, sendo as setas finas relativas à primeira parte (linhas 2 a 6) e as setas grossas relativas à segunda parte contrastada (linhas 7 a 9).

No que se refere aos sinais manuais, o recurso mais evidente de sinalização das partes componentes do contraste são os itens lexicais UM (linha 2) e OUTR@ (linha 7). Enquanto o primeiro sinal dá início à elaboração da primeira parte, o segundo dá início à preparação da segunda parte, contribuindo assim para a sua segmentação. A delimitação da primeira parte é ainda fortalecida pelo item AQUI (linha 6), que explicitamente encerra a primeira parte do contraste antes de R passar à seguinte. Por fim, outro aspecto manual relevante é a espacialização dos sinais. Na primeira parte do contraste, os sinais são realizados à direita do falante, espacialização representada pelo itálico sobre as glosas (linhas 2 a 6); já na segunda parte, a sinalização é realizada à esquerda do falante, espacialização representada pelo negrito sobre as glosas (linhas 7 a 9). A abertura e o encerramento do contraste, por sua vez, são sinalizados de maneira neutra, no espaço de sinalização não-marcado à frente do falante.

A coesão da prática é ainda fortalecida pela grande recorrência lexical e/ou estrutural entre os diferentes componentes do contraste. No que diz respeito à abertura e o fechamento da prática, observa-se uma recorrência não apenas lexical mas sintática, com a frase inicial *DEPENDER PESSOA PESSOA* (linha 1) sendo repetida ao final (linha 10), apenas acrescida de outro sinal *PESSOA*. Com relação às partes contrastantes, internas à prática, a similaridade mais marcante se dá na frase central de cada uma das partes: “*EU PREFERIR ISSO QUÍMICA*”, afirmada na primeira parte, e negada na segunda. Como consequência dessa similaridade lexical e estrutural, as duas partes do contraste diferem uma da outra em apenas 261ms, com a primeira parte apresentando 3s530ms de duração e a segunda, 3s269 de duração. Em conjunto, os recursos segmentais (i.e. itens lexicais de delimitação das partes, recorrência lexical e estrutural) e suprasegmentais (i.e. espacialização dos sinais), relativos aos sinais manuais, permitem não apenas uma identificação da prática como um todo coeso, mas

também a projeção do seu término possível por parte do interlocutor, W, que inicia o seu turno em ótima sincronia com o início da unidade final.¹⁸⁷

A consideração das demais instâncias de contraste do corpus permite a identificação de outros recursos lexicais relevantes para a construção dessa prática na libras – análise que, assim como no caso das listagens, deve ser futuramente aprofundada com uma ampliação do corpus. Dentre os itens lexicais que aparecem já ao início da prática, projetando as suas partes está a o sinal <BÓIA-DOIS>, para o qual a mão ativa direciona o sinal ESS@. Já na fronteira entre a primeira e a segunda parte, além do sinal OUTR@ – que não precisa necessariamente ser precedido do sinal UM ao início da primeira parte – outros recursos que aparecem são os sinais <FICAR> e <DESCULPAR>.¹⁸⁸

Além disso, ainda no que concerne aos recursos manuais, a espacialização dos sinais parece ser um recurso extremamente produtivo na construção de contrastes. A iconicidade na exploração do espaço à frente do sinalizador – com o contraste entre regiões opostas no espaço iconicamente indicando o contraste entre eventos, estados e referentes igualmente antagônicos – é explorada em doze das quatorze instâncias identificadas. Na maioria dos casos, as regiões exploradas são os espaços à [direita/esquerda] do sinalizador,¹⁸⁹ mas foram também identificados oposições [frente/trás] e [interlocutor/falante].¹⁹⁰

¹⁸⁷ Como discutido na análise da figura 29, W inicia o turno em resposta ao contraste de R realizando um extenso fechamento dos olhos e alongamento do sinal EU, o que sugere o aproveitamento de um ponto de relevância para transição (TRP) sem que a sua fala subsequente estivesse inteiramente planejada. Tal ação revela a força do TRP na aproximação do término do contraste, tendo em vista que W busca garantir seu direito de resposta tão logo ela se torna relevante, independentemente de seu despreparo de fato para oferecer uma resposta adequada naquele instante.

¹⁸⁸ Tanto o sinal FICAR quanto o sinal DESCULPAR da libras parecem ter adquirido uma função de conectivo adversativo para o estabelecimento de contrastes, o que revela uma extensão semântica da palavra de conteúdo, com sentido pleno, para uma palavra funcional, com um sentido significativamente esvaziado exceto pela relação contrastiva que estabelece entre a frase que lhe precede e que lhe sucede.

¹⁸⁹ A preferência por essa ordenação (primeiro à direita e depois à esquerda) nos contrastes identificados no corpus *pode* estar relacionada ao fato de os participantes surdos serem destros.

¹⁹⁰ Da mesma maneira que ocorre com os múltiplos recursos de estruturação encontrados na construção das listagens, tais opções não devem ser arbitrárias, de modo que a sua motivação pode e deve ser buscada em futuras investigações. Embora não seja o foco da presente análise, gostaria mais uma vez de tecer observações sobre algumas possíveis motivações desses diferentes recursos.

Na oposição espacial [interlocutor/falante], os sinais referentes a uma das partes do contraste são espacializados na região próxima ao falante, e os sinais referentes à outra parte, na região próxima ao interlocutor. Aqui, a motivação do contraste é clara: o espaço onde cada sinal ou frase é realizado indica,

No que concerne à dimensão não-manual do contraste da tabela 4, vários outros recursos identificados vêm se somar aos recursos manuais para a sinalização de suas partes componentes e da prática como um todo. Um primeiro recurso que aparece no contraste são os breves desvios do olhar, do interlocutor para a região onde as mãos sinalizam, durante a produção da frase de abertura (linha 1). Como mostra a figura 45 abaixo, tais desvios ocorrem simultaneamente à realização dos sinais PESSOA e PESSOA, que são espacializados, respectivamente, à direita e à esquerda de R. Essa observação dá força à hipótese postulada anteriormente, na discussão sobre as soletrações manuais: a de que breves desvios do olhar em meio a um padrão estável salientam a informação sob escopo do olhar divergente. No presente caso, os olhares,

icônica e gestualmente, a qual dos participantes cada uma das partes do contraste se refere. Em <[uma instância do corpus](#)>, por exemplo, R sinaliza a primeira parte do contraste, POBRE, com os braços quase esticados, num espaço de sinalização marcado à frente (i.e. próximo ao interlocutor); em seguida, sinaliza a segunda parte do contraste, EU RIC@, com os braços significativamente dobrados, num espaço de sinalização igualmente marcado (i.e. próximo ao falante). O resultado é um contraste que poderia ser traduzido por algo como: “Você é pobre, mas eu sou rica!”. Esse tipo de oposição [interlocutor/falante], contudo, deve se mostrar significativamente limitado, pelo fato de muitos sinais da libras serem realizados num ponto de articulação fixo no corpo, o que deve limitar a produtividade desse tipo de modulação.

Com relação ao contraste [frente/trás], ele aparece no corpus numa instância de “afirmação/negação” de uma proposição. Na <[instância em questão](#)>, R pergunta para W se, na classe em que ele estudava, havia intérpretes de libras-português. W responde afirmando que havia sim, pois uma “colega” da turma “sabia um pouco de língua de sinais”. Diante dessa resposta, R reage veementemente, produzindo o seguinte contraste:

UM ALUN@ NÃO (trás)	“Não uma aluna ao seu lado!”
SENTAR CORRET@ (frente)	“Alguém devidamente sentado à frente!”

Essa espacialização dos sinais é acompanhada por inclinações do tronco, para trás e para frente. Como mostra Wilbur e Patschke (1998), em sua discussão sobre inclinações do corpo, movimentos para trás na ASL estão associados a algum tipo de *negação* e, movimentos para frente, a algum tipo de *afirmação*. Uma explicação plausível para essa correlação, as autoras sugerem, pode ser buscada nas linhas do trabalho de Lakoff e Johnson (1980), com certas orientações no espaço concreto, como frente e atrás, sendo mapeadas em noções abstratas, como afirmação e negação, por meio de uma rede de metaforização, como por exemplo:

- i) [INTERAÇÃO é um RECIPIENTE];
- ii) [IR PARA ATRÁS é ESTAR FORA] e [IR PARA FRENTE é ESTAR DENTRO];
- iii) [ESTAR FORA é NÃO-PARTICIPAR] e [ESTAR DENTRO é PARTICIPAR];
- iv) [NÃO-PARTICIPAR é NEGAR] e [PARTICIPAR é AFIRMAR].

Por fim, na combinação de contrastes [direita/esquerda e frente/trás], observa-se uma recursividade da prática, com um “subcontraste” sendo contraposto a outro “subcontraste”: o primeiro, estabelecido na oposição frente/trás à *direita*; e o segundo, estabelecido na oposição frente/trás à *esquerda*.

Essas diferentes formas de estruturação sugerem que a segmentação do espaço em regiões delimitadas, embora convencionalizada, seja em última instância um procedimento ad-hoc, realizado de acordo com o tipo de formulação contrastiva que o falante, por razões retóricas específicas, deseja fazer num dado momento.

juntamente com o item lexical PESSOA, estabelecem cataforicamente as duas regiões contrastivas no espaço de sinalização onde cada um dos referentes do contraste serão elaborados, chamando a atenção do interlocutor, já na realização da frase de abertura, para a prática como um todo.



Figura 45. Salientação de itens a serem contrastados por meio de breves desvios do olhar.

Outro recurso não-manual que é utilizado de maneira extremamente produtiva na sinalização das partes contrastantes é a orientação do tronco, que aparece em dez dos quatorze contrastes. Quando utilizado, tal recurso parece ser acompanhado da espacialização dos sinais, embora o inverso nem sempre ocorra (i.e. há instâncias de contrastes que envolvem sinais espacializados, mas não mudanças de orientação do tronco). Em alguns casos, a orientação do tronco para um ou outro lado é acompanhada de um movimento do tronco como um todo para cada um dos lados. Como mostra a figura 46 abaixo, as diferentes orientações do corpo funcionam como gestos atencionais que direcionam o interlocutor com extrema clareza às diferentes partes componentes do contraste, constituindo-se num dos mais eficazes recursos de estruturação da prática.¹⁹¹

¹⁹¹ Apesar da predominância de gestos atencionais tais como a espacialização de sinais manuais e a orientação do tronco na construção dos contrastes, algumas poucas instâncias são produzidas sem qualquer exploração do espaço de sinalização. Nesses casos, o contraste é realizado exclusivamente por meio de mudanças na expressão facial e/ou posição da cabeça, com vários articuladores não-manuais apresentando uma configuração relativamente estável pelo tempo que perdura a primeira parte do contraste, e transformando-se em outra configuração similarmente estável na segunda. Nesses tipos de contraste, não se observa unidades de abertura e fechamento, como na instância em discussão.

Cabe destacar, contudo, que a estabilidade desses sinais não-manuais, principalmente a expressão facial, é apenas relativa. No caso do contraste produzido por R na tabela 4, por exemplo, é possível identificar padrões distintos entre a expressão facial na primeira e na segunda parte do contraste. Como pode ser percebido nas duas imagens que aparecem no centro da figura 46, a primeira parte apresenta uma expressão afetiva de empolgação, enquanto a segunda apresenta uma expressão de desagrado. Essa expressão afetiva corresponde à semântica de cada uma das partes, que envolvem, respectivamente, o *gosto* por química (linhas 2 a 6) e o *desgosto* por química (linhas 7 a 9). Contudo, essa expressão afetiva,

			
Orientação para o interlocutor: Frase disparadora do contraste (linha 1)	Orientação para a direita: Elaboração da primeira parte (linhas 2 a 6)	Orientação para a esquerda: Elaboração da segunda parte (linhas 7 a 9)	Orientação para o interlocutor: Frase de fechamento do contraste (linha 10)

Figura 46. Orientação do corpo segmentando as diferentes partes componentes do contraste.

A complexidade desse contraste, em termos de segmentação da língua, é bem maior do que as instâncias de listagem analisadas na seção 7.4.1. Isso acontece pelo fato de o contraste em questão apresentar apenas dois itens, ambos consideravelmente elaborados em uma série de unidades, diferentemente do que ocorria com as listas, nas quais, predominantemente, cada elemento listado correspondia a uma única unidade. Por esse motivo, uma análise completa sobre a segmentação da prática que aparece na tabela 4 envolve ainda uma reflexão sobre como as unidades internas a cada uma das partes contrastadas (linhas 2 a 6 e 7 a 9) podem ser delimitadas.

Nesse ponto, os achados acumulados nas análises de turnos simples e listas oferecem importante contribuição. Três tipos de alongamentos finais analisados na seção 7.3.1 – representados na tabela 4 pela anotação ‘::’ – contribuem para a delimitação das unidades que pertencem à primeira parte do contraste: i) o sinal UM, formado por uma suspensão independente, tem a suspensão alongada por cinco frames, constituindo sozinho uma única frase (linha 2); ii) o sinal ISSO, também formado por uma suspensão independente, tem a sua fase expressiva transformada num golpe de dois movimentos repetitivos (linha 3); e iii) o sinal PROJETO, formado por um golpe de dois movimentos repetitivos em sua forma de menção, tem esse golpe aumentado para três movimentos repetitivos (linha 5).

No que se refere à aceleração inicial, a principal unidade que apresentou modulações desse tipo foi a que aparece na linha 9, envolvendo modulações nos três primeiros sinais da frase – como mostra a anotação ‘>’ que acompanha as glosas na

claramente perceptível, interage com outras expressões faciais gramaticais que ocorrem ao longo da elaboração de cada uma das partes, o que impede uma estabilidade total dos articuladores faciais.

tabela 4. O sinal *DEPENDER*, por exemplo, é realizado de maneira extremamente abreviada, praticamente se restringindo à movimentação das mãos em relação ao punho; uma execução extremamente simplificada quando comparada à realização plena do sinal, que pode ser observada nas frases 1 e 10. Essa simplificação dos movimentos repetitivos internos ao golpe de um sinal é uma modulação nova nesta investigação, não revelada na análise da seção 7.3.2. As outras duas modulações já haviam sido previamente descritas: a assimilação da configuração de mão de *FAZER* pelo sinal subsequente, *EU* – que também ocorre na linha 5, com a assimilação da configuração do sinal *PROJETO* pelo sinal *EU* – e a redução do número de repetições internas ao golpe no sinal *FAZER*, que passa de duas repetições, em sua forma de menção, para uma no contexto em questão.

Essa unidade apresenta ainda um outro tipo de modulação que, nas seções 7.3.5 e 7.4.1, foi proposto como hipótese a respeito do acento na libras. Representada na tabela 4 pela anotação ‘+’, a modulação do sinal *SIMPLES* envolve uma suspensão pré-golpe de três frames seguida de uma velocidade e tensão do golpe claramente mais intensas do que a dos demais sinais. Esse mesmo tipo de modulação foi encontrado no sinal *PREFERIR*, que aparece na linha 3, e na sua repetição, na linha 8, embora, na segunda instância desse sinal, o acento pareça menos saliente, envolvendo uma suspensão pré-golpe de dois frames e uma velocidade e tensão do golpe mais moderadas. O fato de essas três palavras trazerem informações cruciais para o contraste em questão (i.e. o gosto vs. desgosto por química, e a necessidade de fazer algo “simples” no segundo caso) sugere que a hipótese do acento esteja no caminho certo.

Um outro recurso novo de salientação que as duas partes do contraste revelam é o uso de um outro tipo de bóia: o gesto de apontamento. Liddell (2003) chama esse recurso de *bóia temática*, afirmando que ele é utilizado para salientar porções do discurso de maior relevância, para as quais a bóia aponta. Tal hipótese, perfeitamente compatível com o pressuposto teórico sobre a função atencional do apontamento, é corroborada também pela análise da frase para a qual a bóia temática produzida por R aponta: precisamente as frases das linhas 3 e 8, que, como afirmado acima, trazem as informações cruciais para o estabelecimento do contraste. Como mostra a figura 47 abaixo, a bóia, realizada com a mão esquerda, permanece apontando para o local onde cada uma das frases é sinalizada, somente desaparecendo quando o sinal que envolve as duas mãos, *QUÍMICA*, é realizado.



Figura 47. Utilização de bóia temática para salientação de informações de especial relevância no discurso.

Outro aspecto que contribuiu para a segmentação de várias unidades, também analisado nas seções 7.3.3 e 7.4.1, são os gestos atencionais que co-ocorrem com as frases, tais como expressões faciais e posições da cabeça gramaticais. No contraste em questão, foi possível identificar três gestos desse tipo: i) a inclinação da cabeça para o lado sobre as frases 1 e 10 (ver, e.g., figura 45), que, pelo meu conhecimento intuitivo, funciona pragmaticamente como um modalizador; ii) a forte inclinação da cabeça à frente sobre a frase 4, que, na literatura, tem sido freqüentemente associada à marcação de tópico; e iii) o balanço da cabeça para a esquerda e para a direita de maneira repetitiva sobre a frase 8, indicando negação. A figura 48 abaixo ilustra como é que esse tipo de gesto atencional – no caso, referente ao item (ii) – permite a segmentação de unidades na libras (para análise similar, ver discussão da figura 41, seção 7.4.1).

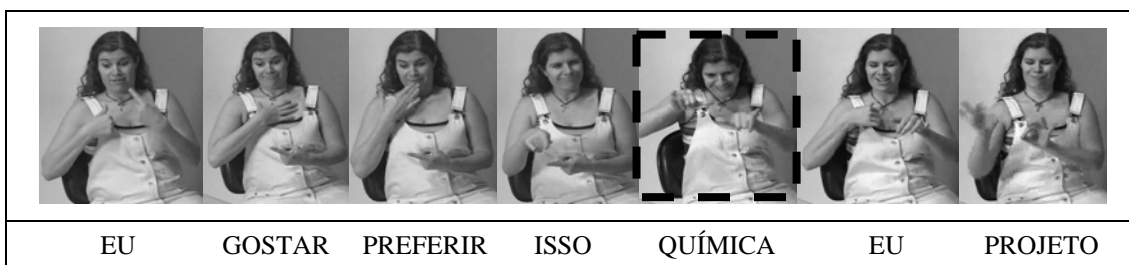


Figura 48. Gesto atencional delimita unidade dentro da primeira parte do contraste.

Um outro recurso não-manual que também se mostrou relevante para a segmentação de unidades em cada uma das partes do contraste foi o restabelecimento do

contato visual ao final das unidades – já apontado por Baker (1977: 223-4). Nesse contraste, em particular, cada um dos itens elaborados envolvia um personagem cuja voz foi incorporada por R. Por essa razão, o padrão não-marcado do olhar na primeira parte do contraste envolvia o acompanhamento da orientação do tronco à direita, e, na segunda parte, à esquerda, e não ao interlocutor. Em meio a esse padrão, então, pôde-se constatar três retomadas do contato visual de R com W (linhas 3, 5 e 9), todas elas em momentos que coincidiam com outras pistas de delimitação de unidades, em especial os alongamentos. Esse recurso de delimitação pode ser visto na figura 49 abaixo, que mostra o padrão do olhar durante a realização da unidade da linha 5 (ver também o término da unidade da linha 3, na figura 48 acima).¹⁹²

¹⁹² A discussão desta seção sobre como os gestos atencionais têm sido utilizados mostra que, enquanto o tronco tende a permanecer mais estável, orientando-se à frente, à esquerda e à direita, a cabeça move-se de maneira mais intensa, ora para acompanhar o próprio tronco, ora para acompanhar o olhar, ora para realizar gestos atencionais que co-ocorrem com a oração (e.g. tópico, negação). O olhar, por sua vez, se revela ainda mais dinâmico, deslocando-se não apenas em consonância com o tronco e a cabeça, mas também ao final das unidades e em pontos discretos em seu interior. Tomados em conjunto, então, esses diferentes gestos mostram uma interação complexa e dinâmica que exige uma explicação.

Interpretando a prática do contraste como um tipo de “objeto” lingüístico que precisa ser atendido, nota-se que o seu formulador tem como uma de suas tarefas sinalizar tanto os limites da prática em seu contexto interacional (suas bordas externas) quanto os limites das diferentes partes que a constituem (as bordas internas, referentes às unidades mínimas e às partes constitutivas da prática). A hipótese é a de que esse trabalho de gerenciamento atencional seja feito precisamente por meio de gestos, que, dependendo do “pedaço” da prática que está sendo sinalizado e do contexto imediato onde ele se insere, interagem de forma hierárquica para estabelecer diferentes níveis de atenção.

Considere, por exemplo, a figura 48. Na imagem correspondente ao sinal ISSO, R re-estabelece o contato visual e reorienta parcialmente a cabeça ao interlocutor, enquanto o tronco permanece voltado para a esquerda. Nesse caso, a combinação hierárquica de gestos atencionais vem indicar, simultaneamente, dois níveis de salientação frente aos quais o interlocutor deve se orientar: um referente à unidade mínima então encerrada (a figura); e outro referente à primeira parte contrastada na prática (o fundo). É como se, por meio da sinalização não-manual em ISSO, R dissesse ao seu interlocutor, *com a cabeça e o olhar*, “Preste atenção, pois acabo de delimitar uma unidade mínima de formulação...”, e *com o tronco*, “...mas não esqueça que essa unidade diz respeito à primeira parte do contraste”. É plausível especular, portanto, que quanto mais complexo for um contraste, maior será a quantidade de recursos formais necessários para a construção de uma prática coesa e eficaz.

Esse emprego do tronco, cabeça e olhos no direcionamento da atenção em mais de um nível revela, tal como argumentam Perrett e Emery (1994) – cf. capítulo 4, seção 4.5 – que o olhar e a cabeça devam ter prioridade atencional sobre o tronco, quando ambos interagem num contexto interacional específico. Tal hierarquia, contudo, como propõem Langton et al. (2000), não implica uma inibição completa dos gestos atencionais hierarquicamente inferiores, mas sim uma sobreposição de camadas com diferentes graus de salientação, ou níveis de atenção, frente aos quais o interlocutor deverá se orientar para gerenciar a sua participação.

A análise sugere ainda que a hierarquia de gestos atencionais deva estar relacionada com a dinamicidade dos articuladores, de forma que, quanto mais dinâmico um dado articulador (i.e. quanto menos grupos musculares ele envolver no seu deslocamento), mais alto/prioritário/importante será o nível de atenção que ele irá sinalizar. No contraste, por exemplo, o tronco e o olhar se colocam,



Figura 49. Retomada do contato visual ao final da unidade.

Por fim, um último recurso não-manual que se mostrou relevante para a segmentação do contraste foi o das *imagens bucais* (ver discussão na nota de rodapé 10 e na seção 5.4). Na unidade que aparece na linha 9, R produz uma imagem bucal, [sɪpl̩s], que perdura pelo tempo que a falante leva para realizar os três últimos sinais da frase, SIMPLES PALAVRA PALAVRA. Esse espraiamento da imagem bucal, do sinal a ela associado (i.e. SIMPLES) para outros sinais subseqüentes, já apontado por Sandler (1999: 22), mostra, na perspectiva cognitivista aqui trabalhada, que o segmento que co-ocorre com a imagem bucal deve ter um estatuto de unidade cognitiva de algum tipo para o falante. De fato, é impossível pensar como três sinais concomitantes a uma única imagem bucal poderiam pertencer a quadros atencionais distintos. Nesse sentido, o espraiamento das imagens bucais pode constituir uma pista para a segmentação de unidades na libras no mesmo sentido que a redução fonética de sinais, isto é, apontando pontos de *junção* na cadeia de fala – em contraposição a marcadores disjuntivos, tais como alongamentos finais, acenos de cabeça, retomada do contato visual e piscadas.

A análise acima mostra que o contraste de R apresentado na tabela 4 foi construído de maneira altamente estruturada. Contudo, assim como ocorre com as listagens, nem todos os contrastes que os falantes procuram expressar são construídos dessa maneira. Embora toda estruturação da fala seja afetada por contingências, menos ou mais explícitas (e.g. a digressão introduzida na lista de W, analisada na figura 41; o novo item acrescentado após o término da lista de R, analisado na figura 43), há momentos em que o processo de construção da fala apresenta um caráter

respectivamente, nos extremos do contínuo “estabilidade \leftrightarrow dinamismo”, com a cabeça situando-se em algum ponto entre os dois. Quando níveis hierárquicos de salientação precisam ser estabelecidos, então, são os articuladores mais dinâmicos que se movem para estabelecer as sucessivas figuras da atenção, enquanto os menos móveis permanecem estáveis para delimitação dos respectivos fundos.

significativamente menos elaborado. Quando isso acontece, no caso dos contrastes, o resultado pode comprometer os propósitos do falante ao empregar tal prática.

Para Atkinson (1984), o sucesso de um contraste pode ser medido pela eficácia com que ele elicia uma resposta coesa do auditório, o que, nos discursos políticos, se dá na forma de aplausos. Essa análise é corroborada não apenas pelos contrastes bem-sucedidos, mas também pelos que fracassam. Nesse último caso, o autor identifica trechos de discurso que, embora expressem contrastes, são construídos de tal maneira que nem todos os participantes são capazes de identificar a prática. Como resultado, os possíveis pontos de completude desses contrastes mal elaborados acabam envolvendo aplausos restritos apenas a uma porção da platéia, ou mesmo nenhum aplauso.

No caso da conversação, em que o(s) interlocutor(es) não assumem a posição de receptores apenas para aclamar o falante-corrente, o que lhes confere autonomia suficiente para concordar e discordar quando lhes parecer conveniente, o sucesso de um contraste não pode ser medido apenas pela ausência ou presença da resposta. Entra em questão, também, se a resposta do interlocutor se mostra afiliativa ou não em relação à posição assumida pelo falante por meio do contraste produzido. Ainda assim, pode-se argumentar que, ao produzir um contraste numa conversa, o objetivo do falante seja o de eliciar uma *resposta positiva*; o que, por sua vez, implica “eliciar uma resposta”, em primeiro lugar, e “uma resposta positiva”, em segundo lugar. A presença ou ausência de resposta, portanto, pode também ser considerada como um critério legítimo do sucesso (parcial) da prática em questão em contextos conversacionais.

Uma das passagens do corpus revela esse jogo sutil entre contingencialidade e estruturação na construção de constrastes, ora ineficazes, ora eficazes. No [<excerto em questão>](#), W estava falando sobre a dificuldade com a apresentação de física que ele teria que fazer perante a turma. W assinala, com espanto, que ele estava responsável pela apresentação de seis conceitos, ao passo que seu colega, Luis (I-L2), fora incumbido de apenas dois, como mostra a *tradução* da passagem abaixo.

- 01 (W) Nós fomos chamados para explicar o tema "Física"
- 02 (W) Eu... (*coça a cabeça em sinal de dificuldade*)
- 03 (W) Eu não conhecia aquelas palavras de significado difícil
- 04 (W) "Por favor, me ajuda"
- 05 (W) A garota sentada ao lado.
- 06 (W) "Aqui, por exemplo, explica" (*aponta para a mão que representa o papel*)

- 07 (W) "Ah, entendi esse, esse, esse item"
- 08 (W) Tinha seis!
- 09 (W) Iten= conceitos.
- 10 (W) Tinha seis!
- 11 (W) Passamos todos eles.
- 12 (W) "Pronto?"
- 13 (W) "Ah, entendi."
- 14 (W) Já o I-L2,
- 15 (W) Do= só dois!
- 16 (W) Só ficou com 2!
- ➔ 17 (W) Eu com seis era muito
- ➔ 18 (W) e ele com dois!
- ➔ 19 (W) Folgado!

Na passagem acima, as duas partes relevantes para a análise estão assinaladas, respectivamente, por flechas pretas (com fundo branco) e brancas (com fundo preto). Ambas as partes envolvem o mesmo contraste (i.e. a quantidade irrisória de conceitos pelos quais Luis ficou responsável na apresentação de física, em contraste com a grande quantidade de conceitos atribuída a W), porém cada um deles é construído de maneira bem distinta. No primeiro contraste (linhas 8 a 16), a despeito da espacialização dos sinais em cada uma das partes, há muito pouca semelhança estrutural entre elas, podendo inclusive notar-se um breve trecho de discurso direto (linhas 12 e 13) entre uma parte contrastada e outra, além de hesitações e reciclagens da fala (linhas 8 a 10 e linha 15 e 16). A análise de Atkinson (1984), assim como a análise do contraste de R nesta seção, contudo, mostram que os contrastes bem sucedidos na eliciação de respostas envolvem uma grande semelhança estrutural, com cada uma das partes apresentando algum tipo de recorrência (lexical e/ou sintática), padrões prosódicos distintos porém coesos, e um tempo de duração relativamente igual.

A má estruturação do primeiro contraste se reflete na total ausência de resposta de R, que não esboça qualquer reação ao seu término na linha 16, como pode ser visto na figura 50 abaixo (35s400ms). Tal ausência possivelmente se deve ao fato de R sequer ter reconhecido o contraste como uma prática em si, diminuindo assim tanto a possibilidade quanto a relevância de uma resposta imediata naquele ponto. De fato, W parece ter sentido a fraqueza do contraste que havia tentado expressar, pois opta imediatamente por reciclá-lo, agora de maneira altamente enxuta e eficaz, como mostra o trecho sob a faixa vertical escura.

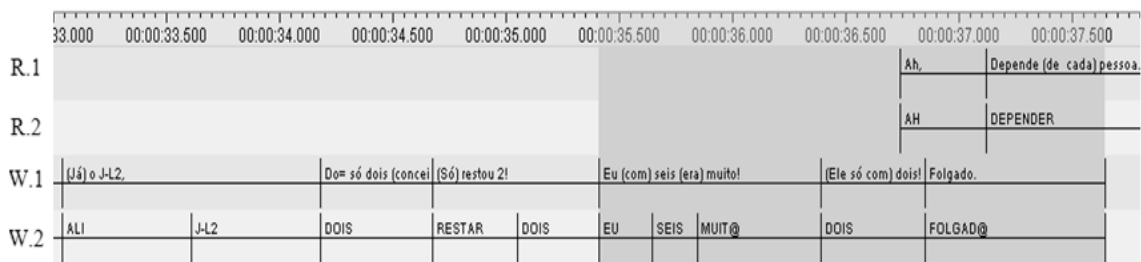


Figura 50. Eficácia na reformulação de um contraste.

O <[contraste reformulado de W](#)>, marcado pelas setas brancas na tradução acima, aparece acompanhado de fotos dos sinais na figura 51. Nele, vários dos recursos metódicos de estruturação dessa prática descritos na presente seção se mostram presentes: i) a espacialização dos sinais, com a primeira parte do contraste (EU SEIS MUIT@) sendo realizada no espaço à frente do falante e a segunda parte (DOIS FOLGAD@), no espaço à esquerda; ii) a mudança de expressão facial, com a primeira parte do contraste envolvendo a boca aberta – ressaltando o espanto com a “grande” quantidade de conceitos – seguida do fechamento da boca na segunda parte; iii) a recorrência estrutural “[expressão referencial] + [predicativo]”, que na primeira parte é realizada, respectivamente, por meio de “[pronome EU + numeral SEIS] + [adjetivo MUIT@]”, ao passo que na segunda parte é realizada por meio de “[espacialização do sinal + numeral DOIS] + [adjetivo FOLGAD@]”;¹⁹³ e iv) a duração aproximada de ambas as partes, que diferem em apenas 300ms uma da outra, a primeira parte apresentando 981ms e a segunda, 1s252ms.

Além disso, um recurso novo observado nesse contraste é a manutenção do sinal SEIS no ar, enquanto o restante da primeira parte e toda a segunda parte do contraste são realizadas. Minha hipótese é a de que esse recurso – uma *bóia de conteúdo*, se é que assim podemos dizer, aproveitando-se do conceito de *bóia* de Liddell (2003) – sirva a dois propósitos: no âmbito da primeira parte, salienta a *informação* relativa ao sinal SEIS, chamando a atenção do interlocutor para a grande quantidade de conceitos pelos quais W ficou responsável (ver, e.g., hipótese de Bolinger, 1961: 25, sobre função atencional do alongamento); já no âmbito da prática como um todo, salienta o *contraste* dessa informação central com a da segunda parte, isto é, o contraste entre “seis” e

¹⁹³ No contexto, a região à esquerda de W já havia sido usada anteriormente para localizar o referente “Luis”. Por causa disso, e devido ao fato de esse referente ter sido estabelecido há pouco tempo nessa região (cf. figura 50, W.2, os sinais ALI I-L2), a simples espacialização do sinal DOIS e FOLGAD@ à esquerda basta para que o referente seja estabelecido.

“dois” conceitos. Esse recurso adicional confere uma saliência e coesão ainda maior ao contraste, como revela a figura 51 abaixo.



Figura 51. Recursos metódicos utilizados na reformulação bem-sucedida de um contraste.

A evidência de que a reformulação do contraste – por meio do emprego de vários recursos metódicos de estruturação dessa prática – se mostrou eficaz pode ser vista na figura 50. Nela, nota-se a pronta reação de R diante do contraste reformulado de W, primeiramente por meio de um gesto típico de inícios de turnos que emerge imediatamente ao término de DOIS, sobrepondo-se a FOLGAD@ (trilha R.2, AH, 36s750ms); e, em seguida, por meio da unidade que dará abertura a um novo contraste em resposta à W, emergindo em perfeita sincronia com a retração total de FOLGAD@ (cf. análise da figura 23).

7.5. Conclusão

O presente capítulo teve como objetivo central elencar um conjunto de pistas formais que contribuam para a identificação de UEs na libras. Ele estende, portanto, a análise sobre segmentação de turnos do capítulo anterior, buscando oferecer critérios para a identificação da organização *interna* de um turno na libras. Tal questão se mostra fundamental não apenas para análises gramaticais dessa língua, mas para a própria compreensão sobre como os falantes surdos coordenam suas participações na conversação.

O capítulo teve início com uma discussão sobre os procedimentos de análise adotados, tendo sido apontada a contribuição da lingüística cognitiva e dos estudos da gestualidade, de um lado, e da AC, de outro. Em conjunto, essas linhas de pesquisa permitiram a elaboração de uma estratégia de análise que envolveu duas abordagens complementares: uma primeira, utilizando o conhecimento sobre as características formais e funcionais das UEs e de gestos envolvidos no processo de gerenciamento da

atenção na interação; e uma segunda, utilizando o conhecimento sobre a organização e função de atividades ou condutas estruturadas da conversação, para as quais as UEs servem como espécies de tijolos de construção.

A escolha metodológica pelos turnos simples, no início da investigação, e dos turnos complexos com níveis progressivos de complexidade, num momento subsequente, foi um procedimento importante para que resultados satisfatórios pudessem ser alcançados. No caso dos turnos simples, compostos por uma única unidade, pode pressupor que quaisquer marcas neles identificadas deveriam estar, de algum modo, relacionadas com a questão da delimitação de UEs. Os achados dessa primeira etapa da investigação, então, facilitaram a segmentação dos turnos complexos estudados em seguida.

Nessa primeira etapa, a análise das pistas formais para a identificação de UEs levantou três tipos principais de recursos: as modulações nos sinais, relativas tanto aos alongamentos finais, presentes nas fronteiras de UEs, quanto às reduções fonéticas, que contribuem para o estabelecimento da coesão interna dessas unidades; os gestos atencionais, formados por expressões faciais e posições de cabeça que co-ocorrem com diferentes tipos de frases; e as piscadas de olhos, que emergem nas fronteiras entre unidades. Além desse levantamento, relativo ao fenômeno da salientação no nível do discurso, uma hipótese sobre a manifestação formal da salientação no âmbito das UEs também pôde ser postulada, com base numa ocorrência em que um sinal fora realizado com uma saliência especial.

Os achados nessa primeira etapa serviram de base, então, para a análise da primeira prática estruturada considerada neste estudo: a listagem. Num primeiro momento, foram considerados aspectos gerais das listas, tal como a sua função na fala-em-curso e a sua organização interna e externa, isto é, em relação ao contexto sequencial imediato. Verificada a compatibilidade em relação às características gerais de listagens nas LOs, passou-se então à análise que mais diretamente interessava à presente pesquisa, voltada para os recursos formais que permitem a sinalização das partes componentes da lista e da prática como um todo.

Nessa análise, alguns recursos previamente identificados, tal como o alongamento final, as reduções fonéticas e a expressão facial que co-ocorre com a frase puderam ser corroborados; além disso, novos recursos também puderam ser

constatados, tais como os movimentos paralelísticos do olhar e/ou da cabeça-tronco que acompanham o paralelismo sintático das frases, e o aceno de cabeça marcado, realizado ao término de UEs que aparecem em pontos de reorientação importantes na organização geral da seqüência de fala.

Por fim, com relação à função atencional do acento, a análise das listas permitiu que a hipótese previamente postulada fosse fortalecida, e uma outra hipótese nova, apresentada. No primeiro caso, a modulação do sinal manual [preparação mais longa + suspensão pré-golpe + golpe mais rápido e intenso], identificada na análise de turnos simples, foi constatada numa frase diferenciando um dentre três sinais fonologicamente idênticos, o sinal modulado estando relacionado justamente à introdução da informação nova no contexto em questão. No segundo caso, os dados permitiram levantar uma outra hipótese sobre a manifestação formal do acento: a de que breve(s) desvio(s) do olhar, de um padrão estável para ponto(s) discreto(s) no curso de uma unidade, resultaria(m) na salientação do conteúdo segmental sob escopo do olhar divergente.

A discussão de uma lista em particular, que envolveu a adição de um novo item após o encerramento da prática, permitiu ainda destacar a primazia dos recursos prosódicos em relação aos recursos lexicais e sintáticos na estruturação dessa prática. Por meio de uma análise sobre quando e como o interlocutor reage diante da produção dessa lista, foi demonstrado que os falantes são capazes de identificar itens seqüencialmente posteriores ao término de uma lista como parte integrante dela; re-análise que é feita com base na percepção de padrões prosódicos paralelísticos próprios da construção das listagens.

A última etapa da análise, então, envolveu a segunda e última prática estruturada que vem sendo considerada nesta tese: o contraste. Assim como na análise das listas, a discussão inicial girou em torno de aspectos gerais relativos aos contrastes, tal como a sua função no discurso em libras e a sua organização interna e externa, na seqüência de fala imediata. Mais uma vez neste caso, foi apontado que os contrastes na libras apresentavam características bastante similares às dessa mesma prática nas LOs, o que sugere, assim como no caso das listas, a sua universalidade.

O passo seguinte foi o de verificar quais são os recursos formais da libras utilizados na sinalização das partes componentes do contraste, bem como da prática como um todo. Nessa análise, destacaram-se como elementos novos a espacialização de

sinais por meio da segmentação do espaço de sinalização em regiões opostas (frente/trás e direita/esquerda, interlocutor/falante) e a orientação do tronco, que, em conjunto, permitiam uma clara visualização das fronteiras internas e externas do contraste.

Contudo, enquanto no caso das listas a identificação das partes componentes implicava (quase) diretamente a identificação de unidades mínimas do discurso, no caso do contraste as duas partes que constituem o seu núcleo, as partes contrastantes, eram, cada uma delas, consideravelmente elaboradas ao longo de várias unidades. Por esse motivo, o conhecimento sobre as UEs mínimas e os gestos atencionais se mostrou mais uma vez fundamental na compreensão da organização interna da prática, possibilitando uma segmentação completa da passagem.

Submetendo cada uma das partes contrastadas a um escrutínio, então, mostraram-se novamente produtivas as pistas relativas ao alongamento final e as expressões faciais e movimentos de cabeça que co-ocorrem com a frase. Além disso, novos recursos também puderam ser identificados, tal como o reestabelecimento do contato visual ao término da unidade, como sinal de disjunção na cadeia de fala, e o espraiamento de imagens bucais para além do sinal à qual elas se referem, como sinal de junção na cadeia de fala. Por fim, no que tange ao acento, a observação de desvios do olhar na unidade de abertura, chamando a atenção para a realização subsequente do contraste, e a observação do mesmo padrão de modulação manual previamente identificado, operando na modificação de sinais cruciais para a semântica das partes contrastadas, vieram fortalecer ainda mais as duas hipóteses previamente postuladas sobre as manifestações formais do acento na libras.

Desse modo, o presente capítulo demonstrou que o procedimento adotado, combinando o conhecimento sobre UEs e gestos atencionais com o conhecimento sobre práticas sociais, constitui uma metodologia eficaz para a descoberta de pistas prosódicas de segmentação do discurso sinalizado em unidades gramaticais. Tal procedimento pode continuar a ser desenvolvido não apenas com a ampliação do corpus, a fim de se aprofundar a análise de listas e contrastes, mas também com a identificação de novas práticas estruturadas (e.g. narrativas) que permitam consolidar a base de conhecimento que começa agora a ser constituída sobre a prosódia da libras.

As marcas formais de salientação da libras no âmbito do discurso, identificadas ao longo dessas três etapas da análise, aparecem sumarizadas na tabela 5 abaixo.¹⁹⁴

Marcas formais prosódicas de segmentação na libras		
Nível	Tipo	Função Prosódica
Manual	<i>Alongamento final</i> a) manutenção da suspensão pós-golpe ou da suspensão independente b) reiteração dos movimentos repetitivos internos ao golpe c) transformação de uma fase expressiva formada por suspensão independente em uma fase expressiva formada por golpe	Marcação de disjunção na cadeia de fala, delimitando fronteiras entre UEs e/ou trechos maiores de discurso
	<i>Reduções fonético-fonológicas</i> a) sobreposição da fase expressiva de um sinal com a fase de preparação de outro sinal b) elisão de movimentos repetitivos internos ao golpe c) retenção bastante breve da suspensão independente nas fases expressivas sem golpe d) assimilação da configuração de mão do sinal subsequente pelo sinal inicial e) abreviação dos movimentos repetitivos internos ao golpe	Marcação de junção na cadeia de fala, fortalecendo a coesão interna da UE
	<i>Gestos atencionais coesos</i> a) espacialização dos sinais	Delimitação de UEs e/ou trechos maiores de discurso
Não-manual	<i>Sinais não-manuais</i> a) piscada de olhos b) acenos de cabeça c) retomada do contato visual	Marcação de disjunção na cadeia de fala, delimitando fronteiras entre UEs e/ou trechos maiores de discurso
	d) espraçamento de imagens bucais	Marcação de junção na cadeia de fala, fortalecendo a coesão interna da UE
	<i>Gestos atencionais coesos</i> a) expressões faciais b) posicionamentos e/ou movimentos da cabeça c) orientações e/ou movimentos do tronco d) direcionamento e/ou movimentos do olhar	Delimitação de UEs e/ou trechos maiores de discurso

Tabela 5. Síntese das marcas formais de salientação da libras no âmbito do discurso.

¹⁹⁴ Considerando o foco central do presente estudo sobre a *prosódia* da libras como meio para chegar às unidades gramaticais, não foram incluídos nesta tabela os recursos lexicais e estruturais identificados ao longo das análises.

Além dessas, as marcas relativas ao fenômeno da salientação no âmbito das UEs aparecem sumarizadas na tabela abaixo:

Marcas formais prosódicas de acentuação na libras		
Nível	Tipo	Função Prosódica
Manual	Modulação das fases do gesto (preparação mais longa, seguida de suspensão pré-golpe e golpe mais rápido e longo)	Salientação de um item informacional numa UE
Não-manual	Breves deslocamentos do olhar, de um padrão estável para pontos discretos no curso de uma UE	Salientação de item informacional e/ou facilitação da percepção de uma soletração manual

Tabela 6. Síntese das marcas formais de salientação da libras no âmbito das UEs.

8. Conclusão

A presente tese vem dar mais um passo no ainda incipiente trabalho de descrição da libras com base em corpus de natureza espontânea ou quase-espontânea. A questão central que norteou o estudo foi a da segmentação do discurso em unidades gramaticais. O objetivo desta conclusão será o de retomar essa proposta de pesquisa, resumindo os principais aspectos tratados; apontando as limitações identificadas; destacando as contribuições teóricas, metodológicas e empíricas; propondo alguns direcionamentos de pesquisa futuros; e refletindo sobre como esse tipo de estudo pode contribuir para a nossa compreensão das línguas naturais de maneira geral.

O argumento que está no cerne desta pesquisa pode ser resumido da seguinte maneira. A fala-em-interação, em especial a conversação, tem um papel central na experiência social humana. É em grande medida por meio dessa atividade que somos capazes de realizar os mais diversos empreendimentos sociais, desde os mais simples até os mais complexos. Ao contrário do que muitos pensam, porém, o estabelecimento de uma conversação envolve um grande investimento *co-laborativo* por parte dos participantes de um encontro. Parte do trabalho conjunto consiste na coordenação da atenção mútua sobre uma mesma entidade a cada novo momento, o que viabiliza não apenas a compreensão apropriada sobre que tipo de ação social está sendo realizada a cada momento, mas também uma organização temporal das participações de cada interlocutor ao longo da interação. No que diz respeito à fala-em-interação e, mais especificamente, à conversação, esse monitoramento e coordenação da ação social envolve, em essência, a produção de recursos tradicionalmente tratados sob o escopo dos estudos da prosódia e da gestualidade. É a gestualidade vocal e corporal, realizada paralelamente ao conteúdo segmental, que oferece as pistas fundamentais aos interlocutores sobre quais pedaços de língua devem ser focados a cada momento, e em relação a qual fundo tais figuras devem ser interpretadas. Essa *salientação* de pedaços de língua, que permite a coordenação da ação social no nível do discurso, é apenas uma das várias manifestações desse fenômeno, que se mostra presente também no âmbito cognitivo, perceptual e interacional da experiência humana.

É com essa tese em perspectiva, então, que os dados da libras foram registrados e analisados em busca de recursos prosódicos/gestuais de salientação empregados no gerenciamento da interação por falantes surdos. No processo de formação do corpus,

procedimentos foram adotados a fim de que se criasse uma situação quase-espontânea de interação nas gravações, de modo a garantir tanto a espontaneidade das conversas quanto a qualidade dos dados a serem analisados. Em seguida, um excerto de uma das conversas foi minuciosamente transcrito, registrando-se o comportamento de diferentes articuladores da libras em diferentes trilhas de anotação num programa de transcrição de dados multimodais, o ELAN.

No processo de análise desse corpus, uma vez demonstrada a relevância do princípio do um-de-cada-vez na conversa em libras, procedimentos foram adotados a fim de viabilizar a busca pelos recursos formais utilizados como pistas de segmentação em cadeias de fala maiores na libras. Um desses procedimentos foi a seleção de instâncias de turnos simples formados por uma única unidade, de modo que as características ali identificadas pudessem, como ponto de partida, ser pressupostas como relevantes para a segmentação do discurso na libras de maneira geral. Outro procedimento foi a seleção de instâncias de turnos complexos formados por listas e contrastes, de modo que a identificação das diferentes partes que constituem essas práticas estruturadas pudessem servir de norte para a corroboração dos recursos formais previamente identificados – na fase de análise anterior –, bem como para a descoberta de novos recursos formais pertinentes à questão da segmentação.

Diante desse trabalho, a principal limitação da pesquisa diz respeito ao caráter restrito do corpus que serviu de base para as análises. Apenas três minutos de uma das conversas em libras foram transcritos de forma exaustiva, e outros excertos das conversas que se mostraram relevantes para a análise (instâncias de listas e contrastes em todas as três conversas gravadas) foram posteriormente consultados e analisados exclusivamente com base na imagem dos vídeos. Tal limitação mostra que, para alcançar uma generalização sólida nas análises, muitas outras investigações, com base em um corpus mais extenso da libras, deverão ser conduzidas.

Esse, porém, é um problema que só poderá ser resolvido com um trabalho intenso de constituição de corpus na libras. Atualmente, estamos num estágio ainda incipiente desse trabalho, que envolve a elaboração de convenções de gravação, transcrição e produção de relatórios. Em especial, enquanto não houver um corpus significativo de dados transcritos, não apenas de conversas, mas de palestras, narrativas, aulas, piadas e inúmeros outros gêneros de fala, toda e qualquer análise sobre aspectos gramaticais da libras terá a generalidade de seus achados comprometida. Tendo em vista

o enorme investimento de tempo que a etapa de transcrição demanda a fim de que todas as informações consideradas relevantes sejam anotadas, acredito que a ampliação do corpus da libras não possa ser realizada senão por meio de um projeto que traga essa meta como um fim em si – ao invés de subordinada a outras metas analíticas, tais como a de identificação de recursos de segmentação do discurso – e que envolva várias pessoas dedicadas aos trabalhos de gravação e, principalmente, transcrição.

No que concerne às contribuições que esta tese oferece, aspectos no âmbito teórico, metodológico e empírico podem ser apontados. No âmbito teórico, destaca-se a proposta de articulação entre o campo da AC e da lingüística interacional, de um lado, e a lingüística funcional e cognitiva e estudos da gestualidade, de outro. A possibilidade e a deseabilidade dessa articulação nasceu de leituras em ambas as áreas, que revelaram, cada uma delas, uma explicação sobre a nossa capacidade de coordenação de ações na interação que, a despeito das perspectivas distintas sobre a função da língua, se mostraram complementares.

Do campo dos estudos da interação, a presente pesquisa trouxe o conhecimento acumulado sobre a relevância da estrutura e da contingência para a projetabilidade dos turnos-na-fala, fenômeno que confere um certo grau de indeterminação ao trabalho de coordenação da participação na interação. A contribuição da pesquisa para essa discussão, então, se deu em duas frentes. Primeiramente, argumentou-se a favor de uma visão de *fala e gramática* que não fique restrita aos recursos lexicais e sintáticos, a fim de que a corporalidade da língua, envolvendo a prosódia e a gestualidade, tenham o seu papel na coordenação da ação social devidamente dimensionado. Em segundo lugar, considerando essa maior variedade de recursos formais que devem entrar na construção dos turnos, foi proposto que os conceitos centrais desse campo (turnos, TCUs e TRPs) recebessem uma definição compatível com a flexibilidade que tais categorias de fato revelam na conversação espontânea. A proposta foi a de que as diferentes formas de exploração dos recursos lexicais, sintáticos, prosódicos e gestuais na construção de unidades da fala deveriam estabelecer, dependendo de sua estruturação, diferentes níveis ou graus de participação na conversação, e, dependendo da convergência ou divergência das projeções estabelecidas por cada recurso isolado, diferentes graus de relevância para transição em pontos discretos ao longo da fala.

Do campo da lingüística funcional e cognitiva, bem como dos estudos da gestualidade, a presente pesquisa trouxe o conhecimento acumulado sobre as

características formais da prosódia e da gestualidade e sobre a sua função básica de salientação, crucial para o gerenciamento da atenção na interação. A contribuição da pesquisa para essa discussão, então, se deu no redimensionamento da relevância dos fenômenos lingüísticos e cognitivos que tais campos descrevem, o gerenciamento da atenção se mostrando relevante não apenas para a troca objetiva de informações, mas principalmente para a negociação de perspectivas e a realização de ações sociais na interação de acordo com os propósitos práticos dos participantes. Essa mudança de perspectiva envolve a necessidade de olharmos para a gramática como inextricavelmente ligada às ações sociais que lhe dão vida, razão pela qual o enunciado em contexto, e não a sentença abstrata, se tornam o objeto de interesse central de pesquisa.

No âmbito metodológico, a contribuição da tese envolve o conhecimento acumulado ao longo desses quatro anos de pesquisa a respeito dos procedimentos envolvidos no processo de formação de corpus: iniciando pela gravação, passando pela informatização dos vídeos e pela transcrição dos dados, até chegar à produção de relatórios a partir das análises. Ainda que um trabalho prévio de nosso grupo de pesquisa tenha sido fundamental para as decisões adotadas neste processo (McCleary e Viotti, 2007), a escolha da conversação como gênero de fala e a necessidade de informatização do corpus fez com que grande parte das decisões tomadas emergissem, de fato, no decorrer desta experiência de pesquisa. Como esperado, tais decisões nem sempre se mostraram as mais adequadas, mas os próprios erros assinalados, ao lado dos acertos, podem servir hoje como ponto de referência para a realização de projetos futuros que envolvam formação de corpus na libras.

No âmbito empírico, então, a contribuição da pesquisa envolve a descrição de fenômenos relativos a diferentes níveis de análise da libras sob uma perspectiva de gramática baseada no uso. Num nível fonético-prosódico, destacam-se as descrições iniciais de diferentes tipos de modulações às quais os sinais são submetidos na fala espontânea, tais como acelerações (e.g. assimilação de traços do sinal subsequente), alongamentos (e.g. reinteração parcial de fases do gesto), intensificações (e.g. combinação de várias modulações para marcação de acento) e espacializações de sinais (e.g. localização marcada da sinalização no espaço à frente do falante).

Ainda nesse nível destaca-se também a proposta de que a função central da orientação do tronco, da cabeça, do olhar e de gestos de apontamento na fala-em-

interação em libras seja, similarmente ao que acontece na interação não-verbal humana e inclusive na animal, a de direcionar a atenção do interlocutor para certas informações em relação a outras. Na libras, argumentou-se que essa função salientadora é também complementada por outros gestos, de função pragmática mais evidente (e.g. vários tipos de expressões faciais e de movimentos de cabeça), que, em conjunto com os demais gestos atencionais, permitem não apenas a delimitação do foco de atenção, mas também o estabelecimento de diferentes níveis de salientação em relação aos quais uma dada figura deve ser interpretada. Por fim, ainda no nível prosódico, destaca-se a descrição de alguns recursos manuais e não-manuais como manifestações do fenômeno do acento na libras.

No nível discursivo, as principais contribuições dizem respeito à descrição inicial das listas e contrastes, unidades lingüísticas ou práticas discursivas com características formais altamente complexas e funções retóricas específicas; e a descrição de recursos metodicamente empregados pelos participantes no gerenciamento da troca de turnos na conversação. No primeiro caso, destaca-se o importante papel coesivo da recorrência lexical, sintática e prosódica na estruturação das listas e contrastes, o que permite ao interlocutor identificar não apenas a organização interna da prática em unidades mínimas e trechos maiores de discurso, mas também a sua organização externa, relativa ao contexto interacional onde ela emerge. No segundo caso, destaca-se o trabalho de projeção e monitoramento fino da fala no processo de troca de turnos, com diversos recursos formais metódicos (e.g. convergência/divergência do olhar, reciclagens da fala, aumento de volume) sendo explorados a fim de se gerenciar, local e colaborativamente, as violações do princípio do um-de-cada-vez na conversação.

As contribuições teóricas, metodológicas e empíricas acima destacadas estão longe de oferecer uma resposta definitiva para as questões de pesquisa que às motivaram. O principal valor desses resultados, a meu ver, está em suscitar uma série de *outras* questões de pesquisa que, se perseguidas, permitirão um maior aprofundamento de nosso conhecimento relativo à descrição da libras, mais especificamente, e à reflexão sobre o funcionamento das línguas naturais, de maneira geral.

Entre as questões que cativam o meu interesse nesses dois âmbitos, podem ser citadas:

i) Dentre os vários gestos atencionais identificados na libras, quais possuiriam uma função essencialmente prosódica, relacionada à estrutura da informação, e quais possuiriam uma função essencialmente pragmática? Ou seria o caso de se questionar, como Bolinger faz, o caráter gramatical da prosódia, entendendo os recursos prosódicos/gestuais como manifestações diretas da emoção humana, com reflexos epifenomenais na gramática?

ii) O quanto o conhecimento acumulado na análise de turnos simples e turnos complexos altamente estruturados pode contribuir para a segmentação do discurso na libras como um todo, inclusive as instâncias de fala mais marcadas pela improvisação e a contingencialidade?

iii) De que modo os vários recursos formais que entram na construção dos turnos da fala interagem, possivelmente de forma hierárquica, para produzir múltiplas trajetórias no turno-em-curso em conversas orais e sinalizadas? É possível corroborar a hipótese de que os pontos de possível completude na fala tenham diferentes graus de relevância para transição, com a trajetória de recursos mais prioritários e/ou de recursos em convergência acarretando trocas mais freqüentes entre falantes do que a trajetória de recursos menos prioritários e/ou de recursos divergentes?

iv) É possível identificar uma orientação dos falantes do português à utilização de gestos de vários tipos, tais como mudanças de orientação do olhar, expressões faciais e gestos emblemáticos? Há alguma maneira de a gestualidade nas LOs ser incorporada no sistema de troca de turnos proposto por SSJ?

v) De que modo os gestos manuais e não-manuais são empregados no gerenciamento da atenção nas LOs? Tais gestos desempenham um papel, por exemplo, na estruturação de listas e contrastes, de maneira similar ao que se observa na libras?

Em todas essas questões, a dimensão gestual emerge como central para as análises, seja na reflexão sobre a gramática da libras, seja na reflexão sobre o papel da corporalidade nas línguas naturais de maneira geral. A meu ver, a análise das LSs, viabilizada pelas tecnologias de registro, manipulação e transcrição de dados multimodais, se mostra particularmente relevante para a lingüística nesse sentido: o de forçar o lingüista a recontextualizar e a “recorporalizar” os dados que lhes servem de base para a análise. Essa *re-situação* da língua no discurso é um projeto que, a despeito de ter sido buscado em campos diversos nas últimas décadas, permanece em certa

medida limitado pela importância que ainda é conferida à escrita (em si, uma grande abstração) *em detrimento da* observação do comportamento situado dos participantes em suas interações cotidianas.

9. Bibliografia

- AMIR, N. et al. (2004). Characteristics of intonation unit boundaries in spontaneous spoken Hebrew: Perception and acoustic correlates. **Speech Prosody Conference**. Nara, Japan, March 23-26, p. 677-680. Disponível em: http://www.isca-speech.org/archive/sp2004/sp04_677.pdf. Acesso em: 15-08-2006.
- ARMSTRONG, D.F.; STOKOE, W.C.; WILCOX, S.E. (1995). **Gesture and the nature of language**. Cambridge: Cambridge University Press.
- ASHBY, W.J.; BENTIVOGLIO, P. (1993). Preferred argument structure in spoken French and Spanish. **Language variation and change**, v. 5, p. 61-76.
- ATKINSON, J.M. (1984). Public speaking and audience responses: Some techniques for inviting applause. In: J.M. Atkinson & J.C. Heritage (Eds.). **Structures of social action: Studies in conversation analysis**. Cambridge: Cambridge University Press, p. 370-409.
- AUER, P. (1996). On the prosody and syntax of turn-continuations. In: Couper-Kuhlen & M. Selting (Eds.). **Prosody in conversation: Interactional studies**. Cambridge: Cambridge University Press.
- AUER, P. (2005). Projection in interaction and projection in grammar. **Text**, v. 25, n. 1, p. 7-36.
- BAKER, C. (1976). What's not on the other hand in American Sign Language. In: **Papers from the Twelfth Regional Meeting of the Chicago Linguistics Society**. University of Chicago.
- BAKER, C. (1977). Regulators and turn-taking in American Sign Language. In: F. Friedman (Ed.). **On the other hand: New perspectives on American Sign Language**. New York: Academic Press, p. 215-241.
- BAKER, C.; PADDEN, C. A. (1978). Focusing on the nonmanual components of American Sign Language. In: P. Siple (Ed.). **Understanding language through sign language research**. New York: Academic Press, p. 27-57.
- BAKHTIN, M. (1979). **Marxismo e filosofia da linguagem**. São Paulo: Editora Hucitec.
- BARKHUYSEN, P.; KRAHMER, E.; SWERTS, M. (2008) The interplay between the auditory and visual modality for end-of-utterance detection. **The journal of the Acoustical Society of America**, v, 123, n. 1, p. 354-365
- BATTISON, R. (1974). Phonological deletion in American Sign Language. **Sign Language Studies**, v. 5, p. 1-19.
- BATTISON, R. (2000). Analyzing signs. In: C. Valli & C. LUCAS (Eds.). **Linguistics of American Sign Language: An introduction**. Washington, DC: Clerc Books: Gallaudet University Press.
- BELLUGI, U.; FISCHER, S. (1972). A comparison of sign language and spoken language. **Cognition**, v. 1, p. 173-200.
- BERGMAN, B.; BOYES-BRAEM, P.; HANKE, T.; PIZZUTO, E. (2001). Sign transcription and database storage of sign information. **Special issue of Sign Language & Linguistics**, v. 4, n. 1-2.

- BICKFORD, A. (2005). Using ELAN. A getting-started guide for use with sign languages. Disponível em: <http://www.und.nodak.edu/dept/linguistics/textbooks/UsingELAN.pdf>. Acesso em: 09-03-2007.
- BILLIG, M. (1987). **Arguing and thinking: A rethorical approach to social psychology**. Cambridge: Cambridge University Press.
- BOLINGER, D.L. (1961). **Generality, gradience, and the all-or-none**. Mouton & Co's-Gravenhage.
- BOLINGER, D. L. (1972). Accent is predictable (if you are a mind-reader). **Language**, v. 48, n. 3, p. 633-644.
- BOLINGER, D. L. (1982). Intonation and its parts. **Language**, v. 58, n. 3, p. 505-533.
- BOLINGER, D. L. (1983). Intonation and gesture. **American Speech**, v. 58, n. 2, p. 156-174.
- BOLINGER, D. (1985). The inherent iconism of intonation. In: J. Haiman (Ed.). **Iconicity in syntax**. Philadelphia, PA: John Benjamins, p. 97-108.
- BOLINGER, D. L. (1998). Intonation in American English. In: D. Hirst & A. Di Cristo (Eds.). **Intonation systems. A survey of twenty languages**. Cambridge: Cambridge University Press, p. 45-55.
- BOYES-BRAEM, P. (1999). Rhythmic temporal patterns in the signing of deaf early and late learners of German Swiss Sign Language. **Language and Speech**, v. 42, n. 2-3, p. 177-208.
- BRAZIL, D. (1985). Phonology: Intonation in discourse. In: T.A. van Dijk (Ed.). **Handbook of discourse analysis: Dimensions of discourse**, v. 2. London: Academic Press.
- BRENTARI, D. (1998). **A prosodic model of sign language phonology**. Cambridge, MA: MIT Press.
- BYBEE, J. (2006). From usage to grammar: The mind's response to repetition. **Language**, v. 82, n. 4, p. 711-733.
- CAPOVILLA, F.C.; RAPHAEL, W.D. (2001). **Dicionário enciclopédico ilustrado trilingüe da língua de sinais brasileira (libras)**, v. I e II. São Paulo: Editora da Universidade de São Paulo/Imprensa Oficial do Estado.
- CAPOVILLA, F.C.; RAPHAEL, W.D. (Orgs.). (2005). **Enciclopédia da língua de sinais brasileira: O mundo do surdo em Libras**, v. 4: Comunicação, religião e eventos. São Paulo, SP: Edusp, Vitae, Feneis, CNPq, Fapesp, p. 457-554.
- CARPENTER, M.; AKHTAR, N.; TOMASELLO, M. (1998). Social cognition, joint attention, and communicative competence from 9 to 15 months of age. **Monographs of the Society for Research in Child Development**, v. 255.
- CHAFE, W. (1980). The deployment of consciousness in the production of a narrative. In: W. Chafe (Ed.). **The pear stories: Cognitive, cultural, and linguistic aspects of narrative production**. Norwood, New Jersey: Ablex Publishing Corporation.
- CHAFE, W. (1987). Cognitive constraints on information flow. In: R.S. Tomlin (Ed.). **Coherence and grounding in discourse**. Amsterdam, Philadelphia: John Benjamins, p. 21-51.

- CHAFE, W. (1988). Linking intonation units in spoken English. In: J. Haiman & S. Thompson (Eds). **Clause combining in grammatical discourse**. Philadelphia: John Benjamins, p. 1-27.
- CHAFE, W. (1994). **Discourse, consciousness, and time**: The flow and displacement of conscious experience in speaking and writing. Chicago and London: The University of Chicago Press.
- CLARK, H.H. (1996). **Using language**. Cambridge: Cambridge University Press.
- CLARK, H.H. (2000). O uso da linguagem. **Cadernos de tradução**, Porto Alegre, n. 9, jan-mar, p. 1-72.
- CLARK, H.H. (2002). Speaking in time. **Speech Communication**, v. 36, p. 5-13.
- CLARK, H.H. (2003). When to start speaking, when to stop, and how. **Proceedings of the ISCA Workshop on Error Handling in Spoken Dialogue Systems (EHSD)**, August 28th-31th, Château d'Oex, Vaud, Switzerland, p. 1-4.
- CLARK, H.H. (2005). Coordinating with each other in a material world. **Discourse and Society**, v. 7, p. 507-525.
- CLARK, H.H.; FOX TREE, J.E. (2002). Using *uh* and *um* in spontaneous speaking. **Cognition**, v. 84, p. 73-111.
- CLARK, H.H.; WASOW, T. (1998). Repeating words in spontaneous speech. **Cognitive Psychology**, v. 37, p. 201-242.
- COATES, J.; SUTTON-SPENCE, R. (2001). Turn-taking patterns in deaf conversation. **Journal of Sociolinguistics**, v. 5, n. 4, p. 507-529.
- COULTER, G. (1982). On the nature of ASL as a monosyllabic language. **Paper presented at the annual meeting of the Linguistic Society of America**. San Diego, CA.
- COULTER, G. (1993). Phrase-level prosody in ASL: Final lengthening and phrasal contours. In: G. Coulter (Ed.). **Phonetics and phonology**. Current issues in ASL phonology, v. 3. San Diego: Academic Press, p. 263-272.
- COULTER, J. (1983). Contingent and a priori structures in sequential analysis. **Human Studies**, v. 6, n. 4, p. 361-376.
- COULTER, J. (1991). Cognition: Cognition in an ethnomethodological mode. In: G. Button (Ed.). **Ethnomethodology and the human sciences**. Cambridge: Cambridge University Press, p. 176-195.
- COUPER-KUHLEN, E.; SELTING, M. (1996). (Eds.). **Prosody in conversation: Interactional studies**. Cambridge: Cambridge University Press.
- CROFT, W. (1995). Intonation units and grammatical structure. **Linguistics**, v. 33, p. 839-882.
- CROFT, W. (2005). Intonation units and grammatical structure in Wardaman and in crosslinguistic perspective. **Australian Journal of Linguistics**. Disponível em: <http://www.unm.edu/~wcroft/Papers/WardamanIUs2.pdf>. Acesso em: 31-01-2007.
- CRUTTENDEN, A. (1986). **Intonation**. Cambridge: Cambridge University Press.

- CRYSTAL, D. (1969). **Prosodic systems and intonation in English**. Cambridge: Cambridge University Press.
- De JONG, K.; ZAWAYDEH, B. (1999). Stress, duration, and intonation in Arabic word-level prosody. **Journal of Phonetics**, v. 32, p. 3-22.
- De MARTINO, J.M. (2005). **Animação facial sincronizada com a fala**: Visemas dependentes do contexto fonético para o português do Brasil. Tese de doutorado, Unicamp.
- DIVELY, V.L. (1998). Conversational repairs in ASL. In: C. Lucas (Ed.). **Pinky extension and eye gaze**: Language use in deaf communities. Washington, DC: Gallaudet University Press, p. 137-169.
- DuBOIS, J.W. (1987). The discourse basis of ergativity. **Language**, v. 63, p. 805-855.
- DuBOIS, J.W.; CUMMING, S.; SCHUETZE-COBURN, S.; PAOLINO, D. (Eds.). (1992). **Discourse transcription**. Santa Barbara: University of California, Santa Barbara.
- DuBOIS, J.; SCHUETZE-COBURN, S.; PAOLINO, D.; CUMMING, S. (1993). Outline of discourse transcription. In: J.A. Edward; M.D. Lampert (Eds.) **Talking data**: Transcription and coding methods for language research. Hillsdale, NJ: Lawrence Erlbaum, p. 45-89.
- DUNCAN, S.J. (1973). Toward a grammar for dyadic conversation. **Semiotica**, v. 9, p. 29-46.
- DUNCAN, S. (2003). Gesture in language: Issues for sign language research. In: K. Emmorey. **Perspectives on classifier constructions in signed languages**. Mahwah: Lawrence Erlbawn Associates, p. 259-268.
- DUNCAN, S. (2005). Gesture in signing: A case study from Taiwan Sign Language. **Language and Linguistics**, v. 6, n. 2, p. 279-318.
- EDELSKY, C. (1981). Who's got the floor? **Language in Society**, v. 10, p. 383-421.
- EDWARDS, D. (1991). Categories are for talking: On the cognitive and discursive bases of categorization. **Theory & Psychology**, v. 1, n. 4, p. 515-542.
- EDWARDS, D. (1998). Discourse and information. Review Essay. **Culture & Psychology**, v. 4, n. 1, p. 91-105.
- EDWARDS, J.; BECKMAN, M.; FLETCHER, J. (1991). The articulatory kinematics of final lengthening, **Journal of the Acoustical Society of America**, v. 89, p. 369-382.
- EGBERT, M.M., (1997). Schisming: The collaborative transformation from a single conversation to multiple conversations, **Research on Language & Social Interaction**, v. 30, p. 1-51.
- EMERY, N. J. (2000). The eyes have it: The neuroethology, function and evolution of social gaze. **Neuroscience and Biobehavioral Reviews**, v. 24, p. 581-604.
Disponível em: <http://www.ece.uvic.ca/~btill/papers/facerec/Emery2000.pdf>. Acesso em: 29-03-2007.
- ENGBERG-PEDERSEN, E. (1993). **Space in Danish Sign Language**. Hamburg: Signum.

- ERICKSON, F. (2004). **Talk and social theory**: Ecologies of speaking and listening in everyday life. Cambridge, Massachusetts: Polity Press.
- FAUCONNIER, G. (1985). **Mental spaces**. Cambridge, Massachusetts: MIT Press.
- FERREIRA-BRITO, L.F. (1984). Similarities and differences in two Brazilian sign languages. **Sign Language Studies**, v. 42, p. 45-56.
- FERREIRA-BRITO, L.F. (1995). **Por uma gramática das línguas de sinais**. Rio de Janeiro, UFRJ: Tempo Brasileiro.
- FORD, C.E. (2004). Contingency and units in interaction. **Discourse Studies**, v. 6, n. 1, p. 27-52.
- FORD, C.E.; FOX, B.A.; THOMPSON, S.A. (1996). Practices in the construction of turns: The TCU revisited. **Pragmatics**, v. 6, p. 427-54.
- FORD, C.E.; THOMPSON, S.A. (1996). Interactional units in conversation: Syntactic, intonational, and pragmatic resources for the management of turns. In: E. Ochs et al. (Eds.). **Grammar and Interaction**. Cambridge: Cambridge University Press, p. 134-184.
- FORD, C.E.; THOMPSON, S.A. (2003). Social interaction and grammar. In: M. Tomasello (Ed.). **The new psychology of language**. vol. 2, Mahwah, NJ: Lawrence Erlbaum, p. 119-43.
- FOX, B.; JASPERSON, R. (1995). A Syntactic exploration of repair in English conversation. In: P.W. Davis (Ed.). **Alternative linguistics**: Descriptive and theoretical modes. Amsterdam: John Benjamins, p. 77-134.
- FOX TREE, J.E.; CLARK, H.H. (1997). Pronouncing 'the' as 'thee' to signal problems in speaking. **Cognition**, v. 62, p. 151-167.
- FRIEDMAN, L. (1975). Space, time, and person reference in American Sign Language. **Language**, v. 51, p. 940-961.
- FRISHBERG, N.; GOUGH, B. (2000 [1973]). Morphology in American Sign Language. **Sign Language & Linguistics**, v. 3, n. 1, p. 103-131.
- GAGO, P.C. (2004). Questões de transcrição em análise da conversação. **Veredas**, UFJF, v. 6, n. 2, p. 89-113.
- GALEMBECK, P.T. (1999). O turno conversacional. In: D. PRETI (Org.). **Análise de textos orais**. Projetos Paralelos (NURC/SP). São Paulo: Humanitas, p. 55-80.
- GARCEZ, P.M. (2001). Deixa eu te contar uma coisa: O trabalho sociológico do narrar na conversa cotidiana. In: B.T. Ribeiro; C.C. Lima & M.T. Lopes Dantas (Org.). **Narrativa, identidade e clínica**. Rio de Janeiro: IPUB-CUCA, p. 83-95.
- GARCEZ, P.M.; LODER, L.L. (2005). Reparo iniciado e levado a cabo pelo outro na conversa cotidiana em português do Brasil. **DELTA**, v. 21, n. 2, p. 279-312.
- GARFINKEL, H. (1963). A conception of, and experiments with, 'trust' as a condition of stable concerted actions. In: O.J. Harvey (Ed.). **Motivation and social interaction**. New York: The Ronald Press, p. 187-238.
- GARFINKEL, H. (1967). **Studies in ethnomethodology**. Englewood Cliffs, NJ: Prentice-Hall.

- GARFINKEL, H.; SACKS, H. (1970). On formal structures of practical actions. In: J. C. McKinney; E. A. Tirvakian (Ed.). **Theoretical sociology: Perspectives and developments**. Appleton-Century Crofts, p. 337-366.
- GIVÓN, T. (Ed.). (1983). **Topic continuity in discourse: A qualitative cross-language study**. Amsterdam: John Benjamins.
- GOFFMAN, E. (1961). **Encounters: Two studies in the sociology of interaction**. Indianapolis: Bobbs-Merrill.
- GOFFMAN, E. (1964). The neglected situation. In: J.J. Gumperz; D.H. Hymes (Eds.). **The ethnography of communication**. Special issue of the *American Anthropologist*, v. 66, n. 6, part 2, p. 133-6.
- GOFFMAN, E. (1967). **Interaction ritual: Essays on face-to-face behavior**. New York: Pantheon Books.
- GOFFMAN, E. (1974). The frame analysis of talk. In: **Frame analysis: An essay on the organization of experience**. Boston: Northeastern University Press.
- GOFFMAN, E. (1981). **Forms of talk**. Philadelphia: University of Pennsylvania Press.
- GOODWIN, C. (1979). The interactive construction of a sentence in natural conversation. In: G. Psathas (Ed.). **Everyday language: Studies in ethnomethodology**. New York: Irvington Publishers: p. 97-121.
- GOODWIN, C. (1981). **Conversational organization: Interaction between speakers and hearers**. New York: Academic Press.
- GOODWIN, C. (1986). Audience diversity, participation and interpretation. **Text**, v. 6, n. 3, p. 283-316.
- GOODWIN, C. (2002). Time in action. **Current Anthropology**, v. 43, p. 1-53.
- GUMPERZ, J.J. (1982). Prosody in conversation. In: **Discourse strategies**. Cambridge: Cambridge University Press, p. 100-129.
- GUMPERZ, J.J.; HYMES, D.H. (Eds.). (1972). **The ethnography of communication**. New York: Holt, Rinehart.
- HOPPER, P. (2001). Grammatical constructions and their discourse origins: Prototype or family resemblance? In: M. Pütz; S. Niemeier; R. Dirven (Eds.). **Applied cognitive linguistics**, v. 1: Theory and language acquisition. Berlin: Mouton de Gruyter, p. 109-129.
- HOPPER, P. THOMPSON, S.A. (1980). Transitivity in grammar and discourse. **Language**, v. 56, n. 2, p.251-299.
- HOUGAARD, A. (2005). Conceptual disintegration and blending in interactional sequences: A discussion of new phenomena, processes vs. products, and methodology. **Journal of Pragmatics**, v. 37, p. 1653-1685.
- IWASAKI, S.; TAO, H. (1993). A comparative study of the structure of the intonation unit in English, Japanese and Mandarin Chinese. **Proceedings of the 9th Annual Meeting of the Linguistic Society of America**, Los Angeles, CA, January.
- IZRE'EL, S. (2005). Intonation units and the structure of spontaneous spoken language: A view from Hebrew. **IDP05 International Symposium on Discourse-Prosody**

- Interfaces.** Aix-en-Provence, France, September, 8th-9th, 20. Disponível em: <http://www.tau.ac.il/humanities/semitic/idp05.pdf>. Acesso em: 15-08-2006.
- JOHNSON, M. (1987). **The body in the mind:** The bodily basis of meaning, imagination and reason. Chicago: The University of Chicago Press.
- HALLIDAY, M.A.K. (1967). **Intonation and grammar in British English.** Mouton: The Hague.
- HANKE, T. (Ed.). (2001). **Interface definitions.** ViSiCAST Deliverable D5-1. Disponível em: <http://visicast.co.uk/members/milestones/ViSiCASTD5-1.pdf>. Acesso em: 14.03.2006.
- HERITAGE, J. (1984). Garfinkel and ethnomethodology. Cambridge: Polity.
- HERITAGE, J.; GREATBATCH, D. (1991). On the institutional character of institutional talk: The case of news interviews. In: D. Boden; D.H. Zimmerman (Eds.). **Talk and social structure.** Cambridge: Polity Press, p. 93-137.
- HOUTKOOP, H.; MAZELAND, H. (1985). Turns and discourse units in everyday conversation. **Journal of Pragmatics**, v. 9, p. 595-619.
- HUTCHINS, E. (2005). Material anchors for conceptual blends. **Journal of Pragmatics**, v. 37, n. 10, p. 1555-1577.
- JEFFERSON, G. (1973). A case of precise timing in ordinary conversation: Overlapped tag-positioned address terms in closing sequences. **Semiotica**, n. 9, p. 47-96.
- JEFFERSON, G. (1978). Sequential aspects of storytelling in conversation. In: J. Schenkein (Ed.). **Studies in the organization of conversational interaction.** New York: Academic Press, p. 219-248.
- JEFFERSON, G. (1984). Notes on some orderlinesses of overlap onset. **Tilburg Papers in Language and Literature**, n. 28, p. 1-28.
- JEFFERSON, G. (1986). Notes on 'latency' in overlap onset. **Human Studies**, v. 9, n. 2-3, p. 153-183.
- JEFFERSON, G. (1990). List construction as a task and a resource. In: G. Psathas (Ed.). **Interactional competence.** Washington, DC: University Press of America, p. 63-92.
- JOHNSTONE, B. (1983). Presentation as proof: The language of arabic rhetoric. **Anthropological Linguistics**, v. 25, p. 47-60.
- KÄRKKÄINEN, E. (1996). Preferred argument structure and subject role in American English conversational discourse. **Journal of Pragmatics**, v. 25, p. 675-701.
- KENDON, A. (1972). Some relationships between body motion and speech. An analysis of an example. In: A. Siegman & B. Pope (Eds.). **Studies in dyadic communication.** Elmsford, New York: Pergamon Press, p. 177-210.
- KENDON, A. (1980). Gesticulation and speech: two aspects of the process of utterance. In: M.R. Key (Ed). **The relationship of verbal and nonverbal communication.** The Hague: Mouton, p. 207-227.
- KITA, S.; van GIJN, I.; van der HULST, H. (1997). Movement phases in signs and co-speech gestures, and their transcription by human coders. **Proceedings of the International Gesture Workshop.** Bielefeld, Germany, September, 17th-19th, p. 23-35.

- KLATT, D. (1976). Linguistic uses of segment duration in English: Acoustic and perceptual evidence. **Journal of the Acoustical Society of America**, v. 59, p. 1208-1221.
- KLIMA, E.; BELLUGI, U. (1979). **The signs of language**. Cambridge, Mass: Harvard University Press.
- KRAHMER, E.; SWERTS, M. (2007). The effects of visual beats on prosodic prominence: Acoustic analyses, auditory perception and visual perception. **Journal of memory and language**, v. 57, n. 3, p. 396-414.
- LABOV, W. (1996). When intuitions fail. In: L. McNair; K. Singer; L. Dolbrin; M. Aucon (Eds.). **Papers from the Parasession on theory and data in linguistics**. Chicago Linguistic Society, v. 32, p. 77-106.
- LAKOFF, G. (1987). **Women, fire and dangerous things**. Chicago: The University of Chicago Press.
- LAKOFF, G.; JOHNSON, M. (1980). **Metaphors we live by**. Chicago: The University of Chicago Press.
- LAMBRECHT, K. (1994). **Information structure and sentence form**: Topic, focus, and the mental representations of discourse referents. Cambridge: Cambridge University Press.
- LANGACKER, R.W. (1986). An introduction to cognitive grammar. **Cognitive Science**, n. 10, p. 1-40.
- LANGACKER, R.W. (1987). **Foundations of cognitive grammar**. Volume 1: Theoretical prerequisites. Stanford, CA: Stanford University Press.
- LANGACKER, R.W. (1991). **Foundations of cognitive grammar**. Volume 2: Descriptive applications. Stanford, CA: Stanford University Press.
- LANGACKER, R.W. (2000). **Grammar and conceptualization**. Berlin & New York: Mouton de Gruyter.
- LANGACKER, R.W. (2001). Discourse in cognitive grammar. **Cognitive Linguistics**, v. 12, n. 2, p. 143-188.
- LANGTON, S.R.H.; BRUCE, V. (1999). Reflexive visual orienting in response to the social attention of others. **Visual Cognit**, v. 6, p. 541-568.
- LANGTON S.R.H.; WATT, R.J.; BRUCE, V. (2000). Do the eyes have it? Cues to the direction of social attention. **Trends in Cognitive Sciences**. v. 4, n. 2, p. 50-59.
Disponível em: <http://arts.uwaterloo.ca/~dsmilek/psych398/langton.pdf>. Acesso em: 29-03-2007.
- LERNER, G.H. (1994). Responsive list construction: A conversational resource for accomplishing multifaceted social action. **Journal of Language and Social Psychology**, v. 13, p. 20-33.
- LERNER, G.H. (1996). On the "semi-permeable" character of grammatical units in conversation: Conditional entry into the turn space of another speaker. In: E.Ochs; E.A. Schegloff & S. Thompson (Eds.). **Interaction and grammar**. Cambridge: Cambridge University Press, p. 238-276.

- LIDDELL, S.K. (1978). Nonmanual signals and relative clauses in American Sign Language. In: P. Siple, P. (Ed.). **Understanding language through sign language research**. New York: Academic Press, p. 59-90.
- LIDDELL, S.K. (1980). **American Sign Language syntax**. The Hague: Mouton.
- LIDDELL, S.K. (1984). THINK and BELIEVE: Sequentiality in American Sign Language. **Language**, v. 60, p. 372-99.
- LIDDELL, S.K. (2003a). **Grammar, gesture and meaning in American Sign Language**. Cambridge: Cambridge University Press.
- LIDDELL, S. K. (2003b). Sources of meaning in ASL classifier predicates. In: K.D. Emmorey (Ed.). **Perspectives on classifier constructions in sign languages**. Mahwah, NJ: Erlbaum p. 199-220.
- LIDDELL, S.K.; JOHNSON, R.E. (1989). American Sign Language: The phonological base. **Sign Language Studies**, v. 64, p. 195-277.
- LINDBLOM, B.; LYBERG, B.; HOLMGREN, K. (1981). **Durational patterns of Swedish phonology: Do they reflect short-term memory processes?** Indiana University Linguistics Club, Bloomington.
- LOCAL, J.; KELLY, J. (1986). Projection and 'silences': Notes on phonetic and conversational structure. **Human Studies**, v. 9, p. 185-204.
- LOCAL, J.; WALKER, G. (2004). Abrupt-joins as a resource for the production of multi-unit, multiaction turns. **Journal of Pragmatics**, v. 36, n. 8, p. 1375–1403.
- LODER, L.L.; GONZALEZ, P.C.; GARCEZ, P.M. (2004). Reparo em terceira posição e intersubjetividade na fala-em-interação em português brasileiro. **Veredas**, UFJF, v. 6, n. 2, p. 117-122.
- LOEHR, D. (2004). **Gesture and intonation**. PhD. Dissertation. Georgetown University. Disponível em: http://www9.georgetown.edu/faculty/loehrd/pubs_files/Loehr04.pdf. Acesso em: 18-02-2008.
- LYNCH, M. (2006). Cognitive activities without cognition? Ethnomethodological investigations of selected 'cognitive' topics. **Discourse Studies**, v. 8, n. 1, p. 95-104.
- MARCUSCHI, L.A. (1986). **Análise da conversação**, São Paulo: Ática.
- MARTINET, A. (1978). **Elementos de lingüística geral**. São Paulo: Martins Fontes.
- MATSUMOTO, K. (2000). Intonation units, clauses and preferred argument structure in conversational Japanese. **Language Sciences**, v. 22, p. 63-86.
- MATSUMOTO, K. (2003). **Intonation units in Japanese conversation: Syntactic, information, and functional structures**. Amsterdam: John Benjamins.
- McCLEARY, L.E. (2003). Technologies of language and the embodied history of the deaf. **Sign Language Studies**, v. 3, n. 2, p. 104-124.
- McCLEARY, L.; VIOTTI, E. (2007). Transcrição de dados de uma língua de sinais: Um estudo piloto da transcrição de narrativas na língua de sinais brasileira. In: H. Salles (Ed.). **Bilingüismo e surdez: Questões lingüísticas e educacionais**. Goiânia, GO: Cênone Editorial.

- McILVENNY, P. (1991). Some thoughts on the study of sign language talk. In: K. Sajavaara et al. (Eds.). **Proceedings of the 3rd Finnish seminar on discourse analysis**. Jyväskylä, p. 187-202. Disponível em: <http://paul-server.hum.aau.dk/paul/research/cv/pubs/sign-talk91.pdf>. Acesso em: 19-03-2005.
- McILVENNY, P. (1995). Seeing conversations: Analyzing sign language talk. In: P. ten Have & G. Psathas (Eds.). **Situated order: Studies in the social organization of talk and embodied activities**. Washington, DC: University Press of America, p. 129-50.
- McNEILL, D. (1992). **Hand and mind**. Chicago: University of Chicago Press.
- MOLDER, H.; POTTER, J. (Eds.). (2005). **Conversation and cognition**. Cambridge: Cambridge University Press.
- MONDADA, L. (2006). Participants' online analysis and multimodal practices: Projecting the end of the turn and the closing of the sequence. **Discourse Studies**, v. 8, n. 1, p. 117-130.
- MOREIRA, R.L. (2007). **Uma descrição da dêixis de pessoa na língua de sinais brasileira: Pronomes pessoais e verbos indicadores**. Dissertação de Mestrado, São Paulo: USP.
- MYERS, S.; HANSEN, B.B. (2007). The origin of vowel length neutralization in final position: Evidence from Finnish Speakers. **Natural Language & Linguistic Theory**, v. 25, p. 157-193.
- NESPOR, M.; SANDLER, W. (1999). Prosody in Israeli Sign Language. **Language and Speech**, v. 42, n. 2-3, p. 143-176. Disponível em: http://sandersignlab.haifa.ac.il/html/html_eng/Nespor&Sandler.pdf Acesso em: 15-08-2006.
- NESPOR, M.; VOGEL, I. (1986). **Prosodic phonology**. Dordrecht: Foris Publication.
- NEWKIRK, D. (1998 [1981]). On the temporal segmentation of movement in American Sign Language. **Sign language and Linguistics**, v. 1, n. 2, p. 173-211.
- OCHS, E.; SCHEGLOFF, E.A.; THOMPSON, S.A. (Eds.). (1996). **Grammar and interaction**. Cambridge: Cambridge University Press.
- ONG, W. (1982). **Orality and literacy: The technologizing of the world**. London and New York: Methuen.
- ONO, T.; THOMPSON, S.A. (1995). What can conversation tell us about syntax? In: P.W. Davis (Ed.). **Alternative linguistics: descriptive and theoretical modes**. Amsterdam: John Benjamins: p. 213-271.
- PARK, J. S.-Y. (2002). Cognitive and interactional motivations for the intonation unit. **Studies in Language**, v. 26, n. 3, p. 637-680.
- PASHLER, H. (1994). Dual-task interference in simple tasks: Data and theory. **Psychological Bulletin**, v. 116, n. 2, p. 220-244.
- PERLMUTTER, D.M. (1990). On the segmental representation of transitional and bidirectional movements in ASL phonology. In: S.D. Fischer; P. Siple (Eds.). **Theoretical issues in sign language research**, v. 1. Chicago: The University of Chicago Press.

- PERRETT, D.I.; EMERY, N.J. (1994). Understanding the intentions of others from visual signals: Neurophysiological evidence. **Cahiers de Psychologie Cognitive**, v. 13, p. 683-694.
- PEZATTI, E. G. (1996). Estrutura argumental e fluxo de informação. In: I.V. Koch (Org.). **Gramática do português falado: Desenvolvimentos**. Campinas: Editora da Unicamp, v. 6, p. 275-299.
- PIERREHUMBERT, J. (1980). The phonology and phonetics of English intonation. Tese de doutorado, MIT. Disponível em: <http://dspace.mit.edu/bitstream/1721.1/16065/1/07492108.pdf>.
- POTTER, J.; WETHERELL, M. (1987). **Discourse and social psychology: Beyond attitudes and behaviour**. London: Sage.
- PRETI, D. (Org.). (1999). **Análise de textos orais**. Projetos Paralelos (NURC/SP). São Paulo: Humanitas.
- QUADROS, R.M. (1999). **Phrase structure of Brazilian Sign Language**. Tese de Doutorado, Porto Alegre, PUC-RS.
- QUADROS, R.M.; KARNOPP, L.B. (2004). **Língua de sinais brasileira: Estudos lingüísticos**. Porto Alegre: ArtMed.
- QUEK, F.; BRYLL, R.; MCNEILL, D.; HARPER, M. (2001). Gestural origo and loci-transitions in natural discourse segmentation. **IEEE Workshop on Cues in Communication**. Kauai Marriot, Hawaii, December, 9, p. 1-8. Disponível em: <http://cobweb.ecn.purdue.edu/~harper/papers/origos.pdf>. Acesso em: 29-03-2007.
- QUEK, F.; BRYLL, R.; MCNEILL, D.; HARPER, M. (2002). Gesture spatialization in natural discourse segmentation. **7th International Conference on Spoken Language Processing**. Denver, CO, September, 16-20, p. 1-4. Disponível em: <http://vislab.cs.vt.edu/Publications/2002/PDFfiles/QueMBH02.pdf>. Acesso em: 29-03-2007.
- REDDY, M.J. (1979). The conduit metaphor: A case of frame conflict in our language about language. In: A. Ortony (Ed.). **Metaphor and Thought**. Cambridge: Cambridge University Press, p. 164-201.
- RIBEIRO, B.T.; GARCEZ, P.M. (Orgs.). (2002). **Sociolingüística interacional**. São Paulo: Loyola.
- ROSCH, E. (1978). Principles of categorization. In: E. Rosch; B. Lloyd (Eds.). **Cognition and categorization**. Hillsdale, NJ: Erlbaum.
- SACKS, H. (1972). On the analyzability of stories by children. In: J.J. Gumperz; D. Hymes (Eds.). **Directions in sociolinguistics: The ethnography of communication**. New York: Holt, Rinehart and Winston, p. 325-345.
- SACKS, H. (1974). An analysis of the course of a joke's telling in conversation. In: R. Bauman & J. Sherzer (Eds.). **Explorations in the ethnography of speaking**. Cambridge: Cambridge University Press, p. 337-353.
- SACKS, H. (1992). **Lectures on conversation**. Vol. II. Oxford: Blackwell.
- SACKS, H.; SCHEGLOFF, E.A.; JEFFERSON, G. (2003). Sistemática elementar para a organização da tomada de turnos para a conversa. **Veredas**, v. 7, n. 1-2, p. 9-73. Tradução de SACKS, H.; SCHEGLOFF, E.A.; JEFFERSON, G. (1974). A simplest

- systematics for the organization of turn-taking for conversation. **Language**, v. 50, n. 4, p. 696-735. Por A.M.S. da Cunha, C.F. Duque, J.R. Medeiros, L.M. Silva, M.P. Borges e M.B.P. Schittini, sob coordenação de M.C.C. Oliveira e P.C. Gago.
- SÁNCHEZ-AYALA, I. (2003). Constructions as resources for interaction: Lists in English and Spanish conversation. **Discourse Studies**, v. 5, n. 3, p. 323-349.
- SANDLER, W. (1989). **Phonological representation of the sign**. Dordrecht: Foris.
- SANDLER, W. (1999). Cliticization and prosodic words in a sign language. In: T. Hall; U. Kleinhenz (Eds.). **Studies on the phonological word**. Amsterdam: John Benjamins. (Current Studies in Linguistic Theory).
- SANDLER, W.; LILLO-MARTIN, D. (2001). Natural sign languages. In: M. Aronoff & J. Rees-Miller (Eds.). **Handbook of linguistics**, p. 533-562. Disponível em: <http://sandersignlab.haifa.ac.il/pdf/Natural%20Sign%20Languages.pdf%20.pdf>. Acesso em: 24-04-08.
- SAPIR, E. (1921). **Language: Introduction to the study of speech**. New York: Harcourt, Brace.
- SAUSSURE, F. de (1970). **Curso de lingüística geral**. São Paulo: Cultrix.
- SCHEGLOFF, E. (1968). Sequencing in conversational openings. **American Anthropologist**, v. 70, n. 6, p. 1075-1095.
- SCHEGLOFF, E.A. (1972). Notes on a conversational practice: Formulating place. In: D. Sudnow (Ed.). **Studies in social interaction**. Glencoe, IL: Free Press, p. 75-119.
- SCHEGLOFF, E.A. (1979). The relevance of repair for syntax-for-conversation. In: T. Givón (Ed.). **Discourse and syntax**. New York: Academic Press, p. 261-286.
- SCHEGLOFF, E.A. (1982). Discourse as an interactional achievement: Some uses of “uh-huh” and other things that come between sentences. In: D. Tannen (Ed.). **Analysing discourse: Text and talk**. Georgetown: Georgetown University Press, p. 71-93.
- SCHEGLOFF, E.A. (1991a). Issues of relevance for discourse analysis: Contingency in action, interaction, and co-participant context. In: E.H. Hovy; D.R. Scott (Eds.). **Computational and conversational discourse**. Berlin: Springer-Verlag, p. 3-35.
- SCHEGLOFF, E.A. (1991b). Conversation analysis and socially shared cognition. In: L.B. Resnick; J.M. Levine; S.D. Teasley (Eds.). **Perspectives on socially shared cognition**. Washington, DC: American Psychological Association, p. 150-171.
- SCHEGLOFF, E.A. (1992). To Searle on conversation: A note in return. In: J.R. Searle et al. (Eds.). **(On) Searle on Conversation**. Amsterdam and Philadelphia: John Benjamins, p. 113-128.
- SCHEGLOFF, E.A. (1996). Turn organization: One intersection of grammar and interaction. In: E. Ochs et al. (Eds.), p. 52-133.
- SCHEGLOFF, E.A. (1997). Whose text? Whose context? **Discourse and Society**, v. 8, n. 2, p. 165-187.
- SCHEGLOFF, E.A. (1998). Reflections on studying prosody in talk-in-interaction. **Language and Speech**, v. 41, n. 3-4, p. 235-263.

- SCHEGLOFF, E.A. (2000). Overlapping talk and the organization of turn-taking for conversation. **Language in Society**, v. 29, n. 1, p. 1-63.
- SCHEGLOFF, E.A. (2002). Accounts of conduct in interaction: Interruption, overlap, and turn-taking. In: J.H. Turner (Ed.). **Handbook of sociological theory**. New York: Kluwer Academic, Plenum Publishers, p. 287-321.
- SCHEGLOFF, E.A. (2007). A tutorial on membership categorization. **Journal of Pragmatics**, v. 39, p. 462-482.
- SCHEGLOFF, E.A.; SACKS, H. (1973). Opening Up Closings. **Semiotica**, v. 8, p. 289-327.
- SCHEGLOFF, E.A.; SACKS, H.; JEFFERSON, G. (1977). The preference for self-correction in the organization of repair in conversation. **Language**, v. 53, n. 2, p. 361-82.
- SCHEMBRI, A.; JONES, C.; BURNHAM, D. (2005). Comparing action gestures and classifier verbs of motion: Evidence from Australian Sign Language, Taiwan Sign Language, and non-signers' gestures without speech. **Journal of Deaf Studies and Deaf Education**, v. 10, n. 3, p. 272-290.
- SCHIFFRIN, D. (1987). **Discourse markers**. Cambridge: Cambridge University Press.
- SCHIFFRIN, D. (1994). Making a list. **Discourse Processes**, v. 17, p. 377-406.
- SCHUETZE-COBURN, S.; SHAPLEY, M.; WEBER, E.G. (1991). Units of intonation in discourse: Acoustic and auditory analyses in contrast. **Language and Speech**, v. 34, p. 207-34.
- SEARLE, J.R. (1992). Conversation. In: J.R. Searle et al. (Eds.). **(On) Searle on conversation**. Amsterdam: John Benjamins, p. 7-29.
- SELTING, M. (2000). The construction of units in conversational talk. **Language in Society**, v. 29, p. 477-517.
- SELTING, Margret. (2003). Lists as embedded structures and the prosody of list construction as an interactional resource. **Interaction and Linguistic Structures (InLiSt)**, n. 35, February. Disponível em: <http://www.uni-potsdam.de/u/inlist/issues/35/inlist35.pdf>. Acesso em: 30-01-2007.
- SELTING, M. (2005). Syntax and prosody as methods for the construction and identification of turn-constructive units in conversation. In: A. Hakulinen & M. Selting (Eds.). **Syntax and Lexis in Conversation: Studies on the use of linguistic resources in talk-in-interaction**. Amsterdam: John Benjamins, p. 17-44.
- SELTING, M.; COUPER-KUHLEN, E. (Eds.). (2001). **Studies in interactional linguistics**. Amsterdam: John Benjamins.
- SELKIRK, E. (1984). **Phonology and syntax: The relation between sound and structure**. Cambridge, MA: MIT Press.
- SILVERMAN, K.; PIERREHUMBERT, J. (1990). The timing of prenuclear high accents in English. In: J. Kingston & M. Beckman (Eds.). **Papers in laboratory phonology**, v. I. Cambridge: Cambridge University Press, p. 72-106.
- SINHA, C. (2005). Blending out of the background: Play, props, and staging in the material world. **Journal of Pragmatics**, v. 37, p. 1537-1554.

- SIPLE, P. (1978). Visual constraints for sign language communication. **Sign Language Studies**, v. 19, p. 95-110.
- SMITH, W. (1996). Spoken narratives and preferred clause structure: Evidence from modern Hebrew discourse. **Studies in Language**, v. 20, p. 163-189.
- STOKOE, W.C. (1960). Sign language structure: An outline of the communication systems of the American deaf. **Studies in Linguistics: Occasional Papers**, v. 8. Buffalo, NY: Department of Anthropology and Linguistics, University of Buffalo.
- STOKOE, W.C. (2001). **Language in hand: Why sign came before speech**. Washington, DC: Gallaudet University Press.
- STOKOE, W.C.; CASTERLINE, D.C.; CRONENBERG, C.G. (1965). **A dictionary of American Sign Language**. Washington DC: Gallaudet College Press.
- STREECK, J. (1995). On projection. In: E. Goody (Ed.). **Social intelligence and interaction: Expression and implication of the social bias in human intelligence**. Cambridge: Cambridge University Press, p. 87-110.
- STREECK, J.; KALLMEYER, W. (2001). Interaction by inscription. **Journal of Pragmatics**, v. 33, p. 465-490.
- SUCHMAN, L. (1987). **Plans and situated actions: The problem of human-machine communication**. Cambridge: Cambridge University Press.
- SUPALLA, T. (2003). Revisiting visual analogy. In: K. Emmorey (Ed.). **Perspectives on classifier constructions in signed languages**. Mahwah: Lawrence Erlbaum Associates, p. 249-257.
- SUPALLA, T.; NEWPORT, E.; SINGLETON, J.; SUPALLA, S.; METLAY, D.; COULTER, G. (s.d.). **The test battery for American Sign Language morphology and syntax**. Manuscrito não publicado e materiais em video. New York: University of Rochester.
- SWERTS, M.; KRAHMER, E. (2006). The importance of different facial areas for signaling visual prominence. **Proceedings of the International Conference on Spoken Language Processing**. Pittsburgh, PA.
- SWISHER, V.; CHRISTIE, K.; MILLER, S. (1989). The reception of signs in peripheral vision by deaf persons. **Sign Language Studies**, v. 63, p. 99-125.
- TALMY, L. (1983). How language structures space. In: H. Pick & L. Acredolo (Eds.). **Spatial Orientation: Theory, Research, and Application**. New York: Plenum Press, p. 225-282.
- TALMY, L. (1996). The windowing of attention in language. In: M. Shibatani & S.A. Thompson (Eds.). **Grammatical constructions: Their form and meaning**. Oxford: Clarendon Press, p. 235-287.
- TALMY, L. (2003). The representation of spatial structure in spoken and signed language. In: K. Emmorey (Ed.). **Perspectives on classifier constructions in sign language**. Mahwah, NJ: Lawrence Erlbaum, p. 169-195.
- TANAKA, H. (2000). Turn-projection in Japanese talk-in-interaction. **Research on Language and Social Interaction**, v. 33, n. 1, p. 1-38.

- TANANGKINGSING, M. (2006). What intonation units can tell us about Cebuano grammar. Disponível em: http://web.ksssp.upd.edu.ph/linguistics/plc2006/papers/FullPapers/II-A-3_Tanangkingsing.pdf. Acesso em: 18-02-2008.
- TAO, H. (1996). **Units in Mandarin conversation**: Prosody, discourse, and grammar. Amsterdam: John Benjamins.
- TAUB, S.F. (2000). Iconicity in American Sign Language: Concrete and metaphorical application. **Spatial Cognition and Computation**, v. 2, p. 31-50.
- TAYLOR, J. R. (2002). **Cognitive Grammar**. Oxford, Oxford University Press.
- te MOLDER, H.; POTTER, J. (Eds.). (2005). **Conversation and cognition**. Cambridge: Cambridge University Press.
- TERAUCHI, M; NAGASHIMA, Y; KANDA, K. (2002). A research on turn-taking in JSL conversation. **Proceedings of International Conference on Information and Technology & Applications (ICITA)**, v. 4, p. 1-4.
- THOMPSON, S.A.; COUPER-KUHLEN, E. (2005). The clause as a locus of grammar and interaction. **Discourse Studies**, v. 7, n. 4-5, p. 481–505.
- TOMASELLO, M. (1992a). **First verbs**: A case study of early grammatical development. Cambridge: Cambridge University Press.
- TOMASELLO, M. (1992b). The social bases of language acquisition. **Social Development**, v. 1, n. 1, 67-86.
- TOMASELLO, M. (1998). Reference: Intending that others jointly attend. **Pragmatics and Cognition**, v. 6, p. 219-234.
- TOMASELLO, M. (2000a). First steps toward a usage-based theory of language acquisition. **Cognitive Linguistics**, v. 11, n. 1-2, p. 61-82.
- TOMASELLO, M. (2000b) Culture and cognitive development. **Current Directions in Psychological Science**, v. 9, n. 2, p.37-40.
- TOMASELLO, M.; FARRAR, M.J. (1986). Joint attention and early language. **Child Development**, v. 57, p. 1454-1463.
- TOMASELLO, M.; HABERL, K. (2003). Understanding attention: 12- and 18-month-olds know what is new for other persons. **Developmental Psychology**, v. 39, n. 5, p. 906-912.
- TOMASELLO, M.; KRUGER, A.C. (1992). Joint attention on actions: Acquiring verbs in ostensive and non-ostensive contexts. **Journal of Child Language**, v. 19, n. 2, p. 311-33.
- TOMASELLO, M.; KRUGER, A.; RATNER, H. (1993). Cultural learning. **Behavioral and Brain Sciences**, v. 16, p. 495-552.
- TOMASELLO, M.; NAMEERA, A.; DODSON, K.; REKAU, L. (1997). Differential productivity in young children's use of nouns and verbs. **Journal of Child Language**, v. 24, p. 373-387.
- TURK, A. (1999). Structural influences on boundary-related lengthening in English. In: J. Ohala; Y. Hasegawa; M. Ohala; D. Granville; A. Bailey (Eds.). **Proceedings of the XIVth International Congress of Phonetic Sciences**. California, Berkeley: Linguistics Department, University of California, p. 237–240.

- UNDERWOOD, G. (1993). Introduction. In: G. Underwood (Ed). **The psychology of attention**, v.1. Aldershot: Elgar, p. xii-xxi.
- Van der KOOIJ, E.; CRASBORN, O.; EMMERICK, W. (2006). Explaining prosodic body leans in Sign Language of the Netherlands: Pragmatics required. **Journal of Pragmatics**, v. 38, p. 1598-1614.
- Van HERREWEGHE, M. (2002). Turn-taking mechanisms and active participation in meetings with deaf and hearing participants in Flanders. In: C. Lucas (Ed.) **Turn-taking, fingerspelling, and contact in signed languages**. Washington, D.C.: Gallaudet University Press.
- WEBER, M. (1978). **Economy and society**: An outline of interpretive sociology. Berkeley and Los Angeles: University of California Press.
- WIENER, M.; DEVOE, S. (1974). **Regulators, channels, and communication disruption**. Research proposal, Clark University.
- WILBUR, R.B. (1994). Eyeblinks and ASL phrase structure. **Sign Language Studies**, v. 84, p. 221-240.
- WILBUR, R.B. (1999). Stress in ASL: Empirical evidence and linguistic issues. **Language and Speech**, v. 42, n. 2-3, p. 229-250.
- WILBUR, R.B. (2000). Phonological and prosodic layering of nonmanuals in American Sign Language. In: K. Emmorey & H. Lane (Eds.). **The signs of language revisited: An anthology to honor Ursula Bellugi and Edward Klima**. Mahwah, NJ: Lawrence Erlbaum Associates, p. 215-244.
- WILBUR, R.B.; MARTINEZ, A.M. (2002). Physical correlates of prosodic structure in American Sign Language. **Paper presented at the Meeting of the Chicago Linguistics Society**, April 25-27.
- WILBUR, R.B.; NOLEN, S.B. (1986). The duration of syllables in American Sign Language. **Language and Speech**, v. 29, n. 3, p. 263-80.
- WILBUR, R. B.; PATSCHKE, C.G. (1998). Body leans and the marking of contrast in American Sign Language. **Journal of Pragmatics**, v. 30, p. 275-303.
- WILCOX, S. (2004). Cognitive iconicity: Conceptual spaces, meaning, and gesture in signed languages. **Cognitive Linguistics**, v. 15, n. 2, p. 119-147.
- WINSTON, E.A. (2000). It just doesn't look like ASL! Defining, recognizing, and teaching prosody in ASL. **Proceedings of the 13th National Convention of CIT**. Silver Spring, MD: RID Publications, p. 103-116.
- XAVIER, A. N. (2006) **Descrição fonético-fonológica dos sinais da língua de sinais brasileira**. Dissertação de mestrado. São Paulo, USP.