



Veredas atemática

Volume 20 nº 2 – 2016

Um sistema para transcrever a fala-em-interação: GAT 2¹

Margret Selting, Peter Auer, Dagmar Barth-Weingarten, Jörg Bergmann, Pia Bergmann, Karin Birkner, Elizabeth Couper-Kuhlen, Arnulf Deppermann, Peter Gilles, Susanne Günthner, Martin Hartung, Friederike Kern, Christine Mertzlufft, Christian Meyer, Miriam Morek, Frank Oberzaucher, Jörg Peters, Uta Quasthoff, Wilfried Schütte, Anja Stukenbrock, Susanne Uhmann²

Tradução e adaptação para o português (NUCOI/UFMG): Ulrike Schröder³, Mariana Carneiro Mendes, Caroline Caputo Pires, Diogo Henrique Alves da Silva, Thiago da Cunha Nascimento e Flavia Fidelis de Paula

Revisão técnica: Paulo Cortes Gago (UFJF/UFRJ)

¹ O artigo original foi publicado no seguinte periódico: Gesprächsforschung – Online-Zeitschrift zur verbalen Interaktion (ISSN 1617-1837) Ausgabe 12 (2011), Seite 1-51 (www.gespraechsforschung-ozs.de)

² Gostaríamos de agradecer a Thomas Schmidt por suas contribuições às discussões que levaram a este texto [agradecimentos do texto original].

³ Eu/Ulrike Schröder agradeço/agradece ao Conselho Nacional de Desenvolvimento Científico e Tecnológico – CNPq pela Bolsa de Produtividade (2015-2018), bem como a FAPEMIG pelo apoio que recebo pelo Programa Pesquisador Mineiro (2015-2017).

Abstract

This article presents a revised version of GAT, a transcription system first developed by a group of German conversation analysts and interactional linguists in 1998. GAT tries to follow as many principles and conventions as possible of the Jefferson-style transcription used in Conversation Analysis, yet proposes some conventions which are more compatible with linguistic and phonetic analyses of spoken language, especially for the representation of prosody in talk-in-interaction. After ten years of use by researchers in conversation and discourse analysis, the original GAT has been revised, against the background of past experience and in light of new necessities for the transcription of corpora arising from technological advances and methodological developments over recent years. The present text makes GAT accessible for the Portuguese-speaking community. It presents the GAT 2 transcription system with all its conventions and gives detailed instructions on how to transcribe spoken interaction at three levels of delicacy: minimal, basic and fine. In addition, it briefly introduces some tools that may be helpful for the user: the German online tutorial GAT-TO and the transcription editing software FOLKER.

Keywords: Conversation analysis, transcription, prosody.

Resumo em português

Este artigo apresenta a tradução portuguesa da versão revisada do GAT, um sistema de transcrição, primeiramente, desenvolvido por um grupo de pesquisadores alemães no campo da Análise da Conversa (AC) e Linguística Interacional (LI) em 1998. O GAT busca seguir tantos princípios e convenções quanto possível do sistema de transcrição formulado por Jefferson para o campo da AC, embora proponha algumas convenções mais compatíveis com análises linguísticas e fonéticas da língua falada, especialmente para a representação da prosódia na fala-em-interação. Depois de dez anos de uso por pesquisadores em AC e Análise do Discurso (AD),⁴ o GAT original foi revisado perante a experiência prática e devido a novas necessidades para a transcrição de corpora emergentes de avanços tecnológicos e desenvolvimentos metodológicos durante os últimos anos. A presente tradução torna o GAT acessível à comunidade de falantes da língua portuguesa. Apresenta o sistema da transcrição GAT 2 com todas as suas convenções e dá instruções detalhadas sobre como transcrever uma interação falada em três níveis de precisão: mínimo, básico e refinado. Ademais, introduz, de forma sucinta, algumas ferramentas úteis para o usuário: o tutorial em alemão do GAT-TO, disponível online, e o software FOLKER para editar transcrições.

Palavras-chave: Análise da Conversa, Análise do Discurso, Transcrição, Prosódia.

1. Introdução

⁴ Nota dos tradutores: O termo aqui utilizado refere-se à tradição anglo-americana de Análise do Discurso (*Discourse Analysis*), e não à “L’ecole Française d’analyse du discours”, representada por pesquisadores como Pêcheux, Ducrot, Maingueneau e Foucault.

2. Princípios de transcrição
3. Convenções de transcrição (GAT 2)
 - 3.1. Partes da transcrição
 - 3.1.1. Cabeçalho da transcrição
 - 3.1.2. Estrutura geral da transcrição
 - 3.2. A transcrição mínima
 - 3.2.1. Segmentos e ordenação das palavras
 - 3.2.2. Estrutura sequencial
 - 3.2.3. Outras convenções de transcrição segmental
 - 3.2.4. Mais convenções para a transcrição mínima
 - 3.3. A transcrição básica
 - 3.3.1. Prosódia
 - 3.3.2. Convenções adicionais para a transcrição básica
 - 3.4. A transcrição refinada
4. Sugestões para a representação de fenômenos mais complexos
 - 4.1. Representação autossegmental da entonação
 - 4.2. Notação de ritmo
5. Traduções interlineares
6. Sumário das convenções de transcrição GAT 2
7. Exemplo de transcrição do excerto de uma conversa autêntica
8. Referências

Apêndice 1: Tutorial em alemão do GAT-TO disponível online.

Apêndice 2: Editor de transcrição FOLKER.

1. Introdução

O GAT – cujo acrônimo significa *Gesprächsanalytisches Transkriptionssystem* (sistema de transcrição para análise da conversa) – é um sistema de transcrição para notação da fala e prosódia da fala-em-interação cotidiana autêntica. Pode ser usado tanto para a compilação de transcrições da fala-em-interação a serem trabalhadas com propósitos de pesquisa quanto para transcrições em publicações no campo da linguística.

Uma primeira versão do GAT, inicialmente elaborada para o alemão, foi publicada há dez anos (cf. SELTING et al., 1998). Nesse ínterim, uma série de publicações no campo da AC e AD em países e comunidades de pesquisa falantes de alemão adaptaram esse sistema para transcrição de trechos de dados. Isto sugere que o objetivo original de estabelecer um sistema unificado para a transcrição da interação falada tem sido atingido, pelo menos entre os pesquisadores falantes do alemão. Dez anos depois, desenvolvimentos recentes no campo da prosódia e multimodalidade, bem como a divulgação da LI, novas possibilidades computacionais de gravação e armazenamento de dados e, finalmente, muitos anos de experiência no uso de GAT tornaram aconselhável a adaptação do GAT original (GAT 1) às novas exigências. O GAT 2 (SELTING et al., 2009) fornece uma versão revisada do GAT 1, que esclarece ambiguidades e faz pequenas modificações em falhas que se tornaram evidentes em seu uso diário. Compreende todas as convenções necessárias para a produção de transcrições publicáveis e pode ser lido e entendido

sem conhecimento do sistema anterior. O presente artigo é uma tradução e adaptação de Selting et al. (2009), que visa a tornar o GAT 2 acessível para a comunidade de falantes de português.

Quais as vantagens do GAT 2?

- Pode ser usado, especificamente, para pesquisas nos campos de AC, AD e LI;
- É facilmente acessível para principiantes em transcrição;
- De início, oferece um nível simples de transcrição: a transcrição mínima. A transcrição mínima é suficiente para notar, de maneira simples, o discurso oral, tal como é feito frequentemente, por exemplo, na sociologia ou na psicologia;
- É apropriado para transcrição com programas comuns de processamento de textos, nos quais o *output* é semelhante ao *input* tanto quanto possível. Para transcrever com editores especializados em transcrição, as convenções precisariam, talvez, ser modificadas ou determinadas de forma mais precisa. O GAT 2 tem o objetivo de ser compatível com convenções de transcrição para editores especializados (ele define um *output* possível), mas não busca definir tais convenções.
- A transcrição mínima é, particularmente, compatível com corpora eletrônicos e sustentáveis de transcrição, isto é, com corpora independentes da plataforma, bem como com eficientes algoritmos de busca computadorizada.⁵ Ela permite a notação exata de respiração, pausa e prolongamento. Fenômenos interacionais como riso, choro e sobreposição são representados de forma inequívoca, e são formalizáveis com editores de transcrição e outras ferramentas de processamento.
- O sistema inclui sugestões para a notação de fenômenos mais complexos em linhas separadas. Para a descrição detalhada de fenômenos entonacionais, são discutidas representações autosssegmentais. Há convenções de notação para ritmo e instruções explícitas para traduções interlineares.

O GAT 2 reconhece o fato de que a análise de aspectos visuais em interação multimodal está se desenvolvendo rapidamente. Nessa área, tanto teoria quanto tecnologia estão mudando constantemente, e ainda não há convenções estabelecidas. Além disso, a complexidade do campo vai além do que é manejável no espaço limitado desse artigo. Por isso, sugestões para a notação de aspectos visuais em interação multimodal foram excluídas aqui. Para leitores interessados, convém salientar o trabalho recentemente publicado por Charles Goodwin (e.g. 2007a, 2007b, 2010), Lorenza Mondada (e.g. 2007a, 2007b, 2008) e Jürgen Streeck (e.g. 2003, 2009) (cf. também HEATH, HINDSMARSH & LUFF, 2010).

⁵ “Eficiente” no sentido de evitar tanto sub-seleção (i.e. negligenciando instâncias relevantes) como sobre-seleção (i.e. incluindo instâncias irrelevantes) quanto tanto possível (cf., por exemplo, EDWARDS, 1995).

Além disso, ferramentas adicionais foram desenvolvidas, as quais são descritas detalhadamente no apêndice deste artigo:

- Com base no GAT 2, um tutorial online (GAT-TO) está disponível agora em alemão. O tutorial introduz aos principiantes a prática de transcrição em alemão e lhes fornece módulos de aprendizagem online voltados para aspectos desafiadores da transcrição com o GAT (<http://paul.igl.uni-freiburg.de/GAT-TO/>; cf. também a descrição no apêndice 1).
- Para transcrever com GAT, Thomas Schmidt (Hamburg) desenvolveu o editor de transcrição FOLKER no Instituto para a Língua Alemã (*Institut für Deutsche Sprache – IDS, Mannheim*), disponível na versão atual uma interface de usuário em alemão, inglês ou francês, respectivamente. Esse editor confere automaticamente a validade da transcrição em relação às convenções de uma transcrição mínima no formato GAT. FOLKER foi introduzido para facilitar a compilação de um Corpus de Alemão Falado para Fins de Pesquisa e Ensino (*Forschungs- und Lehrkorpus gesprochenes Deutsch – FOLK*). O editor pode ser baixado gratuitamente no site do Arquivo de Alemão Falado (*Archiv für Gesprochenes Deutsch – AGD*) no IDS (<http://agd.ids-mannheim.de>; cf. também a descrição no apêndice 2). Também pode ser usado para a transcrição de dados de áudio e como base para importação de dados para o ELAN, o software mais comumente usado para análise de vídeo.

Na era da internet, o GAT 2 seria inimaginável sem seu próprio site e links para as ferramentas mencionadas acima (GAT-TO e FOLKER). O URL é o seguinte:

http://agd.ids-mannheim.de/gat_en.shtml⁶

Por fim, convém destacar que uma transcrição sempre depende dos objetivos específicos, e potencialmente variáveis, de análise e descrição. Embora o GAT possa ser usado como fio condutor para transcrição em todos os níveis de granularidade, da transcrição em fase de trabalho à versão refinada com particularidades elaboradas, sua intenção é, antes de tudo, estabelecer padrões mínimos para publicações linguísticas. As convenções de notação para esta finalidade são delineadas sob o título de *transcrição básica*. A depender dos propósitos específicos do pesquisador, a *transcrição básica* pode ser reduzida ao nível da *transcrição mínima* (para discussões de dados internos, por exemplo), ou expandida para o nível da *transcrição refinada*, ou mesmo ir além disso, para publicações com objetivos especiais.

2. Princípios de transcrição

O GAT foi desenvolvido com base nos seguintes princípios:⁷

⁶ Neste link, tem-se acesso a uma versão do tutorial em inglês. No site do IDS (*Institut für Deutsche Sprache*) há também uma versão em alemão.

- Expansibilidade da notação para maior granularidade (“casca de cebola” ou princípio das camadas múltiplas): uma transcrição com baixa granularidade deve ser expansível sem revisão para um maior grau de granularidade.
- Legibilidade da transcrição: a transcrição deve ser compreensível para linguistas e leigos. Isso implica em não usar qualquer modo especial de representação para a fala, como, por exemplo, o alfabeto fonético. Este poderia certamente ser acrescentado para fins específicos de pesquisa.
- Clareza: o sistema de transcrição é destinado a ser explícito ao fornecer a representação de fenômenos auditivos no nível específico de granularidade. Símbolos de transcrição são definidos de modo inequívoco; cada símbolo representa um (e somente um) fenômeno.
- Iconicidade: Pretende-se que os símbolos de transcrição sejam o menos arbitrário possível. Por exemplo: eles devem seguir princípios icônicos.
- Relevância: O sistema de transcrição deve permitir a notação de tais fenômenos, cujas pesquisas anteriores tenham mostrado ser relevantes para a interpretação e análise de interação verbal.
- Parametrização baseada na forma: As convenções de notação aspiram por transcrições baseadas na forma. Isto é, ao invés de comentários interpretativos, como, por exemplo, “surpreso”, os parâmetros específicos nos quais tais interpretações são baseadas devem ser representadas individualmente com base em sua forma.

3. Convenções de transcrição (GAT 2)

3.1 Partes da transcrição

3.1.1. Cabeçalho da transcrição

Quando uma transcrição ou parte dela é citada em uma publicação, ela deve ser precedida de um cabeçalho da transcrição contendo os seguintes metadados:

- fonte do *corpus* e da gravação;
- indicação do início e fim do excerto;

⁷ Ver as contribuições em Edwards e Lampert (1993), que apresentam uma seleção de sistemas de transcrição para AD em uso na época. Para sugestões mais recentes, conferir Jefferson (2004), Schegloff (2007, p. 265-269) e Du Bois (2006).

- uma breve descrição do contexto interacional entre parênteses duplo (()) onde for necessário.

Às vezes, transcrições são simplificadas pelo bem do argumento em discussão (por exemplo, quando em conversas quebradas/divididas, a parte irrelevante poderia ser omitida). Este tipo de simplificação deve ser mencionado no cabeçalho da transcrição.

Se partes da transcrição original estiverem omitidas quando integradas numa publicação, isso deve ser sinalizado por ((...)). Maiores omissões podem ser indicadas da seguinte maneira:

.
· ((20 segundos omitidos))
.

3.1.2. Estrutura geral da transcrição

Uma transcrição do GAT representa as ações linguísticas e não linguísticas, bem como eventos, em sua ordem temporal, em registros de áudio e vídeo. Tais ações e eventos podem ser ligados a referências temporais específicas nos registros, dando marcações de tempo absolutas. Estas provêm horas, minutos e segundos, por exemplo, {1:01:04} para 1 hora, 1 minuto e 4 segundos. Opcionalmente, tal marcação de tempo pode ser fornecida em uma coluna à esquerda de cada linha de transcrição.

Ao escolher uma fonte equidistante como a Courier 10 pt, tal como foi feito para os exemplos deste texto, problemas fruto de conversões entre sistemas operacionais serão minimizados. O espaçamento entre as linhas deve ser normalmente de 1.5. O uso de *tabs* deve ser evitado; em vez disso, o número apropriado de espaços deve ser inserido. Separar palavras por hífen não é permitido; de forma geral, não há hifenização no GAT 2. A transcrição inteira é escrita em letras minúsculas. (Letras maiúsculas são necessárias para a notação de acentos.)

Na transcrição corrente, como na escrita padrão, a sequência de itens na tela ou no papel é uma reflexão icônica da sequência temporal de eventos em tempo real. A direção habitual de leitura da esquerda para a direita e de cima para baixo captura – com algumas poucas exceções – a sequência temporal linear da fala-em-interação.

A transcrição é subdividida em segmentos. Os segmentos são enumerados.⁸ Se a necessidade de espaço para a notação de um segmento for superior ao que é oferecido por uma única linha, o segmento continuará na linha seguinte. Este, contudo, é normalmente recuado e não lhe é atribuído um número. O número do segmento é seguido (após três espaços) pela identificação do falante.⁹ Identificações de falantes não são repetidas no segmento seguinte quando ele mesmo prossegue. Após mais três espaços, a notação da própria elocução continua. A

⁸ Excertos de transcrições mais longas reproduzidas em uma publicação começam ou com o número do segmento relevante da transcrição original ou com o número do segmento 01.

⁹ Para falantes não identificáveis, identificações como X1, X2,..., ou F1, F2, ... e M1, M2,... para falantes femininos e masculinos, respectivamente, podem ser usadas.

transcrição básica pode ser estendida acrescentando linhas extras abaixo do texto, por exemplo, de modo a dar uma representação mais detalhada da prosódia e/ou fenômenos visuais e para traduções para outra língua (estas linhas adicionais não têm numeração; ver também seção 5). Uma seta ‘→’ à esquerda da linha pode ser usada para destacar um fenômeno relevante para a análise. O excerto (1), abaixo, ilustra a estrutura geral de uma transcrição.¹⁰ Outras convenções serão explicadas na seção 3.2.

Exemplo (1): ((conversa fictícia, 1:04-2:05 sec))

```
{1:04} 01 A: aqui é onde a transcrição começa
        02 B: sim exatamente
        03      (--)
        04 A: se você me inter[romper ]
→       05 B:                [não estou] te
        interrompendo
{2:00} 06 A: sim cê está
        07      (.)
        08      cê acabou de fazer isso
        09      acabou de (.) me interromper
        10 B: okay
{2:05} 11      me desculpa
```

Uma margem ampla à direita (pelo menos 5 cm, ou 2 in) possibilitará ao transcritor expandir a transcrição mínima a uma transcrição básica (ver seção 3.3).

3.2 A Transcrição mínima

A transcrição mínima é a forma mais simples de representação de discurso com o GAT. Embora seja suficiente para vários propósitos nas Ciências Sociais (tal como análise de conteúdo em entrevistas), normalmente não é detalhada o bastante para propósitos de pesquisas na Análise da Conversa e na Linguística Interacional. No contexto da Linguística Interacional, a transcrição mínima é principalmente uma ferramenta de trabalho, e não se destina à publicação. Pode ser usada, por exemplo, para determinar as passagens que serão notadas mais detalhadamente. O exemplo 1 (ver acima) apresenta tal transcrição mínima.

Na transcrição mínima, são notadas as elocuções dos participantes e sua segmentação em unidades entonacionais (ver seções 3.2.1, 3.2.2 e 3.3.1). A transcrição mínima também inclui uma

¹⁰ A seguir, trabalharemos primeiramente com exemplos fictícios para ilustrar os fenômenos que gostaríamos de introduzir. Um exemplo de uma conversa cotidiana autêntica será então apresentado no fim deste artigo.

indicação de sobreposições, marcadores de hesitações, pausas, inspirações e expirações, risos, ações e eventos não linguísticos, bem como trechos de fala que são ininteligíveis (ver seções 3.2.2 - 3.2.4). Os finais de segmentos são marcados por quebras de linhas.

Em áreas de pesquisas não linguísticas, transcrições mínimas podem ser também usadas em publicações. Ao representar passagens de monólogos mais longos, como, por exemplo, em entrevistas de Ciências Sociais, pode ser vantajoso marcar os limites dos segmentos por uma barra vertical (U+007C)¹¹ e numerar as linhas consecutivamente para economizar espaço. Deve-se ter em mente, entretanto, que este método torna mais difícil a expansão da transcrição em uma fase posterior. O exemplo 2 mostra a notação contínua dos segmentos.

Exemplo (2):

```
01 INT: e quando foi a última vez que você teve esse problema |
02 NAR: bem | ahm | (0.35) | foi aproximadamente | não sei | você
03 realmente | se você quer saber a hora exata que aconteceu
04 | (0.50) | eu diria provavelmente | deixe-me ver | (1.23) |
05 deve ter sido há treze meses | ou mais |
```

3.2.1 Segmentos e ordenação das palavras¹²

Adaptação ortográfica

Em geral, a notação de uma elocução é simbolizada de acordo com as convenções ortográficas de uma dada língua, as quais são basicamente regras para a conversão dos seus segmentos sonoros para a escrita. Geralmente, a variedade padrão nacional de uma língua é tomada como norma, a menos que uma variedade regional seja escolhida pelos participantes e explicitamente introduzida dessa forma no cabeçalho da transcrição ou em um comentário. Em situações em que existem muitas variedades de padrão nacional, a variedade adotada pelos participantes deve ser especificada no cabeçalho da transcrição. Variações de pronúncia dentro da norma adotada são notadas de acordo com a ortografia padrão, ou seja, elas não são representadas separadamente. Apenas as que fogem da norma adotada são capturadas como fugas da ortografia padrão.¹³ No português brasileiro,¹⁴ isso implica que, por exemplo, a redução do gerúndio em

¹¹ Colchetes com “U+” e combinações de letras e cifras seguindo um sinal indicam a codificação de caracteres no padrão Unicode.

¹² Nota dos tradutores: A partir desta seção, as partes do texto em que se tratou de fenômenos próprios da língua inglesa foram adaptadas e/ou acrescidas pelos tradutores para o português brasileiro, incluindo os exemplos e as referências bibliográficas relacionados aos ditos fenômenos, mediante autorização dos autores originais.

¹³ Quando são necessários mais detalhes fonéticos em uma transcrição, isso pode ser notado em uma linha separada, por exemplo, de acordo com as convenções do AFI, mas veja também a seção sobre ortografia modificada abaixo.

algumas variantes da língua (com a supressão do /d/ e a pronúncia do /o/ átono como /u/) deveria ser grafada como a norma padrão:

/leno/	[leno]	= lendo
	[lendo]	= lendo

No entanto, se a transcrição segue uma variante do PB em que esse tipo de pronúncia não costuma ocorrer, então tal desvio pode ser notado:

/lendo/	[lendo]	= lendo
	[leno]	= leno

Cliticização

Nas cliticizações, a palavra cliticizada perde sua integridade fonética. Geralmente, cliticizações envolvem variados processos concomitantes de redução, tais como supressões e/ou assimilações e/ou reduções vogais e consoantes completas. Frequentemente, tais cliticizações ocorrem com combinações de palavras funcionais, como por exemplo, pronomes clíticos juntos ao verbo (e.g.: pronome átono + verbo). De acordo com a norma padrão da língua, que se baseia no sistema de colocação pronominal do português europeu (SIMÕES, 2010), o pronome átono em posição enclítica é ligado ao verbo por um hífen, no entanto, não pode ser notado desta forma em uma transcrição de acordo com o GAT, uma vez que este símbolo é utilizado neste sistema para a notação de entonação final. Em português brasileiro, o clítico preferencialmente antecede o verbo (SIMÕES, 2010; VIEIRA, 2005), sendo assim, a cliticização não é notada graficamente. Menos convencionalizadas na ortografia, mas ainda observáveis, são as combinações que podem ser cliticizadas.

D'água – De água

Pra – Para a

Num – Em um

Nos casos de dúvidas em que a cliticização particular é convencionalizada, dicionários online podem ser consultados. Dependendo do interesse do pesquisador e do propósito da transcrição, vários outros níveis de granularidade de palavras nas elocuições são possíveis no GAT¹⁵. Se um desses é escolhido, deve ser notado no cabeçalho da transcrição.

¹⁴ Para uma comparação do fenômeno do apagamento do /d/ no morfema de gerúndio no dialeto de São José do Rio Preto, São Paulo, e uma revisão de estudos deste fenômeno em diferentes dialetos do português brasileiro, ver Ferreira (2010).

¹⁵ Esta e a seguinte subseção se afastam da versão em língua inglesa do GAT 2, porque os tradutores sentem a necessidade de adaptar esta seção para público falante de português, suas necessidades e históricos.

Ortografia padrão modificada

Para transcritores de todas as línguas em que há uma diferença significativa entre a ortografia padrão e a pronúncia casual, uma decisão deve ser tomada quanto à maneira de se representar as realizações fonéticas concretas. Também pode ser desejável para fins argumentativos explicitar o tipo específico de língua falada ou captar a maneira como ‘soam’ as palavras. Nestes casos, pode ser usada a ortografia padrão modificada, também chamada de “dialeto visual”. O “dialeto visual” traz a vantagem de estar mais de acordo com a realidade da língua falada, mas a desvantagem de (i) ser difícil de decifrar para falantes não nativos, (ii) possivelmente caricaturar os falantes como tipos sociais e (iii) comprometer a precisão de buscas automáticas.

Para as transcrições aqui apresentadas, os tradutores do GAT 2, após cuidadosa consideração, optaram por uma forma moderada e sistemática da ortografia padrão modificada. Uma aplicação abrangente desta ortografia pode ser encontrada nos exemplos de transcrição de *Du ou Sie*¹⁶ na seção 7 deste artigo. Ao decidir quando e como modificar a ortografia padrão nestas transcrições, fomos guiados pelos seguintes princípios:

- Utilize a ortografia padrão quando forem aplicados processos fonéticos e fonológicos comuns (obrigatórios):

anos – não: *anus* (linha 08)¹⁷

tamanho – não: *tamanhu* (linha 03)

estranho – não: *istranhu* (linha 44)

- Indique variantes fonéticas e fonológicas não regulares (opcionais), modificando a grafia de tal forma que as realizações concretas sejam reconhecíveis:¹⁸

tá – *está* (linha 23)

cê – *você* (linha 23)

- Para apagamentos no começo de uma palavra, utilize um apóstrofo para representar a omissão:

'tava – *estava* (linha 51)

'ocê – *você* (linha 26)

¹⁶ Esse foi o nome dado ao trecho transcrito ao final deste artigo, pois nesse trecho os participantes falam sobre suas experiências em relação ao uso dos pronomes de tratamento formal, *Sie* (senhor/senhora), e informal, *Du* (tu/você), da língua alemã.

¹⁷ Aqui, os números de linha se referem às transcrições da seção 7.

¹⁸ Cliticizações podem ser notadas dependendo do grau de redução, que varia de univerbação, apostrofação, ligação por meio de subscrito ‘_’ à separação simples. Escrever com subscrito ao invés de usar univerbação ou apostrofação melhora a possibilidade de busca automática por itens clíticos.

- Para apagamentos no fim de uma palavra, acentue a vogal anterior para representar a omissão:

perguntô – perguntou (linha 15)
fazê – fazer (linha 82)

- Para apagamentos de sons e sílabas no meio da palavra, deixe de fora a(s) letra(s) correspondente(s):

comprano – comprando
falano – falando

- As modificações ortográficas, no entanto, tentam preservar a possibilidade de identificação das palavras em questão.

Para todas as modificações, é recomendado o uso de um conjunto controlado de formas, e uma lista com as formas utilizadas deveria acompanhar cada transcrição. Esta lista poderia incluir formas reduzidas frequentemente utilizadas na língua falada em sua pronúncia quase padronizada:

num, né, véi, fizero, entrano

Variações de notação de pronúncia

Variantes de notação fonética mais detalhadas podem ser usadas, se um dos propósitos da transcrição é destacar certo aspecto fonético. Para ilustração, veja os possíveis níveis de granularidade para os segmentos 22 e 23 do exemplo de transcrição da seção 7:

a) Ortografia padrão

18 aí o diogo olhou assim e falou
19 e falou que que você está achando velho

b) Ortografia padrão modificada (moderada)

18 aí o diogo olhou assim e falou
19 e falou que que cê tá achano véi

c) “Dialeto visual”

18 aí u diogu olhô assim e falô

Até mesmo níveis mais refinados utilizariam parcial ou integralmente os símbolos do Alfabeto Fonético Internacional (AFI) para representar a fala completa.

Formas dialetais regionais

Quando expressões regionais são relevantes para o problema de pesquisa, uma linha com a transcrição fonética pode ser acrescentada abaixo da respectiva linha transcrita. (De forma alternativa, a variedade regional utilizada por um ou mais falantes pode ser indicada em um comentário no cabeçalho da transcrição.)

Palavras em língua estrangeira

Palavras em língua estrangeira são escritas conforme prescrito pela ortografia da língua original. Por exemplo:

gestalt
*angst*¹⁹

No entanto, quando a pronúncia diverge notavelmente do padrão, isto é notado. Por exemplo, de acordo com a pronúncia, nota-se *gemütlich* (vogal posterior fechada arredondada) ao invés de *gemütllich* (vogal anterior fechada arredondada).

Palavras compostas

No GAT, compostos separados por hífen (*arco-íris*, *couve-flor*) são apresentados sem hífen, já que ele é necessário para a notação de entonação (ver 3.3.1):

arcoíris
couveflor

Abreviações

Quando um falante utiliza letras, abreviações ou acrônimos, no GAT, eles são escritos por extenso, para que os acentos possam ser notados onde for preciso. Então:

u éfe ême gê – UFMG
i bê gê é – IBGE

Números

¹⁹ A ortografia alemã exige que os substantivos sejam iniciados por letra maiúscula; no entanto, isto é evitado no GAT, em que letras maiúsculas são reservadas para a indicação de ênfase em acento primário e secundário.

Números também são transcritos por extenso, por exemplo,

vinte dois
cento e dez
um ponto três
um e meio
meio dia e vinte

Outros

Extensões do alfabeto que não são utilizadas na língua em questão não são notadas na transcrição: *fiica, fantoma* (não *fiică, fantomă*); *sot, dusman* (não *soț, dușman*). Caso necessário, um comentário pode ser fornecido, tal como <<pronúncia romena>> > (ver 3.2.3).

3.2.2. Estrutura sequencial

Contribuições do falante

A contribuição de um falante inclui todos os eventos linguísticos e não linguísticos que seguem imediatamente um ao outro, e que são produzidos por este falante. Uma contribuição de um falante pode ser composta de um ou mais segmentos. O GAT separa cada segmento em uma linha. Um segmento pode consistir de:

- uma unidade entonacional²⁰ (ver seção 3.3.1.), incluindo atraso inicial ou interno ao segmento (e.g.: ‘gaguejamento’);
- uma unidade entonacional e elementos prosodicamente dependentes dela (incrementados ou prefaciados) (ver seção 3.2.3.);
- uma unidade entonacional incompleta;
- uma pausa;
- uma ação corporal visível.

Pausas e eventos que não podem ser atribuídos a algum falante específico também são notados em uma linha separada.

Sobreposições e fala simultânea

[] Colchetes abertos são inseridos no ponto exato da fala em que começa a
[] sobreposição e colchetes fechados, no lugar em que ela termina. Os respectivos
colchetes são alinhados uns com os outros.

²⁰ O termo ‘unidade de fraseação’ (*Phrasierungseinheit*) utilizado no GAT 1 foi substituído pelo termo ‘frase entonacional’ no GAT 2.

Exemplo (1'): ((conversa fictícia, 1:04-2:00 seg))

01 A: é aqui que a transcrição começa
02 B: é exatamente
03 (--)
→ 04 A: se você me interrom[per]
→ 05 B: [eu] não estou interrompendo você
(...)
07 A: você acabou de fazer
08 você (.) me interrompeu

O excerto (1') começa com pequenas contribuições de A e B, que consistem em uma única unidade entonacional. Entre a contribuição de B, na linha 2, e de A, na linha 4, segue-se uma pausa, que não pode ser atribuída a um falante específico. Não pertence nem ao turno de A, nem ao de B, e é, portanto, notada como um segmento separado (linha 3). Na linha 5, B inicia um novo turno em sobreposição a A, que ainda não havia chegado ao fim do seu turno. Nas linhas 8 e 9, A produz um turno que consiste de duas unidades entonacionais, com a segunda apresentando uma micropausa interna.

Depois de cada sobreposição, novas contribuições do falante sobreposto/interrompido são inseridas como um novo segmento. As únicas exceções são continuadores e vocalizações curtas, que não correspondem a uma reivindicação de turno. Nesse caso, um leitor, cujos olhos estejam se movendo da esquerda para a direita e de cima para baixo, deve brevemente retornar de um segmento posterior para um segmento anterior na transcrição:

01 A: eu só [que]ria [fa]lar
02 B: [hm] [hm]

Quando vários segmentos com sobreposições se seguem uns aos outros, o raro problema de identificação dos respectivos pares de sobreposição pode ocorrer. No presente caso, os segmentos que pertencem um ao outro podem ser marcados pelo menor espaçamento entrelinhas (espaçamento simples em vez de 1,5):

01 A: [eu só queria dizer] outra coisa
02 B: [eu também quero dizer algo]
03 A: [quem vai primeiro]
04 B: [eu sempre tenho que] brigar pelo turno

Ou índices podem ser utilizados nos colchetes:

- 01 A: ₁[eu só queria dizer] ₁ outra coisa
 02 B: ₁[eu também quero dizer algo] ₁
 03 A: ₂[quem vai primeiro] ₂
 04 B: ₂[eu sempre tenho que] ₂ brigar pelo turno

Especialmente nos casos mais complicados, deixar linhas vazias entre os grupos dos colchetes indexados pode conduzir a uma melhor compreensão.

Inspiração e expiração

Inspirações e expirações audíveis são notadas com um “h” (n vezes) e um sinal de grau ° (U+00B0). Para inspirações, a marcação com o grau deve preceder o “h” e para expirações, suceder:

- °h / h° in-/expiração audível de aproximadamente 0,2-0,5 seg de duração
 °hh / hh° in-/expiração audível de aproximadamente 0,5-0,8 seg de duração
 °hhh / hhh° in-/expiração audível de aproximadamente 0,8-1,0 seg de duração

Se a respiração durar por mais tempo, essa duração é indicada em um comentário, por exemplo: “((expiração de 2 seg))”.

Pausas

Pausas curtas podem ser ou medidas ou estimadas.

- (.) micro pausa estimada até aproximadamente 0,2 seg de duração
 (-) pausa curta estimada em aproximadamente 0,2 – 0,5 seg de duração
 (--) pausa intermediária estimada em aproximadamente 0,5 – 0,8 seg de duração
 (---) pausa longa estimada em aproximadamente 0,8 – 1,0 seg de duração
 (0.4) mensuração da pausa de 0,4 segundos

Pausas longas são sempre indicadas em segundos (notação até o décimo de segundo).

- (2.0) } Pausa mensurada ou estimada de 2,0 e 2,3 segundos,
 (2.3) } respectivamente

Pausas dentro de uma unidade entonacional são notadas no próprio segmento. Quando uma pausa pode ser inequivocamente atribuída a um dos participantes, por exemplo, antes do início de turno de um próximo falante selecionado, ela é notada como parte do segmento do respectivo falante. Veja o exemplo a seguir:

Exemplo (3): ((conversa fictícia, 1:04-3:00 seg))

01 A: é aqui que a transcrição começa
02 B: é exatamente
03 (--)
04 A: se você me interrom[per]
05 B: [eu] não estou interrompendo você
06 A: sim você está
07 (.)
08 A: você acabou de fazer
09 você (.) me interrompeu
10 B: oh
11 eu fiz mesmo
12 A: (.) fez
13 B: certo
14 me desculpe
15 A: deixa pra lá

Uma pausa dentro de uma unidade entonacional pode ser percebida na linha 9. Aqui, a unidade entonacional iniciada antes da pausa teve continuidade após ela.

A linha 12 apresenta uma pausa no início do turno que pode ser atribuída a um único falante. Da mesma forma, pausas produzidas na extensão do turno de um falante em particular como, por exemplo, durante a projeção das narrativas, são colocadas no respectivo segmento transcrito (e não em linha separada): pausas separando as unidades entonacionais são colocadas ao final do segmento precedente, e pausas de hesitação, no começo do segmento seguinte.

As pausas nas linhas 3 e 7 não podem ser inequivocadamente atribuídas a algum falante específico. Portanto, elas são notadas como segmentos separados. Quando tal pausa não imputável ocorre, o próximo segmento deve ser atribuído novamente a um dos falantes, ainda que esse seja o falante que possuía o turno antes da pausa, como ocorre na linha 8.

3.2.3. Outras convenções de transcrição segmental

Marcadores de hesitação

eh, ãhm, ùhm etc.

marcadores de hesitação, chamados de “pausas preenchidas”²¹

²¹ A representação desses marcadores precisa ser adaptada às especificidades de cada língua. Deve-se ainda tomar cuidado para evitar homografia com itens lexicais.

Risos e choros

hahaha }
hehe }
hihi }

risada “silábica” curta, de acordo com a qualidade vocal e o número de pulsos e sílabas, respectivamente²²

((ri)), ((chora))

caracterização de evento não linguístico

<<riendo> então> }
<<chorando> então> }

fala entremeada com riso ou choro é notada por meio de uma descrição dentro de um conjunto de parênteses angulares (ver 3.2.4.). A duplicação do parêntese angular indica a extensão do fenômeno.²³

<<:-) > então>

“smile voice” (voz de sorriso)

Continuadores

hm, mm }
sim, é, não }
hm_hm, ãhãm }

partículas monossilábicas

partículas bissilábicas

?hm?hm

partícula reduplicada, normalmente utilizada para negação (? = U+0294 (a respeito do uso de ?, ver seção 3.3.2.))

Inícios de turno não acentuados e adjunções (tags)

Inícios de turno não acentuados²⁴ como, por exemplo, *eh*, *ah*, *mas* e adjunções não acentuadas, tais como *sabe*, *então*, *né*, *não são*, usualmente, unidades entonacionais independentes, mas sim elementos prosódicos dependentes. Em outras palavras, eles se anexam, ou se juntam, à unidade entonacional completa seguinte ou precedente, respectivamente (para detalhes, conferir 3.3.1.). Quando isso ocorre, eles são notados na mesma linha que a respectiva unidade entonacional e conectados a ela pelo sinal de igual ‘=’ (U+003D).

x: eh=isso vai funcionar=né

²² Há ainda muito pouco conhecimento acerca do *status* de sílabas proeminentes e menos proeminentes no riso. Se relevantes, partículas de riso mais proeminentes podem ser notadas como acentos, isto é, com letras maiúsculas.

²³ A inserção de partículas de riso dentro de palavras deve ser evitada, como em *ass(h)im*, uma vez que isso torna buscas automáticas mais difíceis.

²⁴ ‘Inícios’ é usado aqui para incluir o que análises mais precisas talvez revelem ser exemplos de marcadores discursivos de pré-inícios, inícios ou prefaciadores de turno (cf. SCHEGLOFF, 1996).

3.2.4. Outras convenções para a transcrição mínima

Ações e eventos vocais não verbais

São notados somente ações e eventos vocais não verbais que são relevantes para a interação.

((funga)), ((tosse)) caracterização de ações e eventos vocais não verbais

((tosse, 10 seg)) ação ou evento com indicação de sua duração

Ações e eventos vocais não verbais podem ocorrer dentro de um turno ou no lugar de um ou mais turnos. Por exemplo:

01 A: eu não quis dizer ((funga)) isso assim
02 B: ((suspira))
03 (.)
04 B: mas foi o que você falou

Eles também podem acompanhar ações verbais:

01 A: eu não quis <<chorando> dizer isso assim>
02 B: ((suspira))
03 (.)
04 B: <<levantando> mas foi o que> você falou

Descrições e comentários em parênteses angulares, como no caso de <<chorando> > e <<levantando> >, devem ser geralmente usados da seguinte maneira: os parênteses angulares interiores separam o descritor, ou o comentário, da transcrição do turno. Os parênteses exteriores indicam seu escopo. O descritor/comentário é alocado naquele ponto da transcrição, em que o fenômeno se torna observável. O parêntese angular mais exterior é alocado naquele ponto da transcrição, em que o fenômeno relevante deixa de ser observável.

Eventos extralinguísticos que não podem ser atribuídos a nenhum falante específico são notados em uma linha sem qualquer identificação de falante, por exemplo, como em ((telefone toca)), ((gravador toca)). Se necessário, eventos simultâneos podem ser notados de maneira semelhante às sobreposições. Eventos que são atribuídos a vários participantes são notados da mesma maneira: ((Joana e Cláudio dão um aperto de mãos)). Em contrapartida, uma ação atribuída a um participante específico é notada como um segmento com a devida identificação do mesmo participante.

Quando ações e eventos vocais não verbais ocorrem simultaneamente, a duração da ação vocal ou do evento não verbal pode ser dividida, a fim de indicar precisamente quando cada evento começa/termina. Por exemplo:

04 B: <<levantando> mas foi o que> você falou
05 A: ((suspira 2 seg)) [((suspira 1 seg))]
06 B: [você acha que eu] posso compensar isso?

Inteligibilidade

()	trecho incompreensível
(xxx xxx xxx)	trecho incompreensível com indicação da duração em sílabas, sendo que cada 'xxx' representa uma sílaba
(posso)	termo presumido, cuja transcrição é incerta
(posso/deixa_eu falar)	sons incertos, ou sílabas, a nível lexical, com possíveis alternativas
((incompreensível, aprox. 3 seg))	trecho incompreensível com indicação de sua duração (notada como ações e eventos vocais não verbais)

Teoricamente, incertezas/alternativas podem, além disso, ser aplicadas a várias palavras em uma sequência. No caso, essas são notadas da mesma maneira: (*traga o Humberto para o Brasil/traga um bom para o Brasil*).

3.3 A transcrição básica

No nível da transcrição básica, a transcrição mínima é expandida para incluir informações prosódicas, o que é necessário para evitar interpretações erradas dos segmentos no seu contexto interacional (em termos da estrutura semântica e função pragmática). Esta seção introduzirá as noções de unidade entonacional, acento focal e movimento entonacional no final do segmento. Ademais, apresentaremos as convenções de notação para outros fenômenos, como fala engatada, prolongamento, fechamento glotal e comentários interpretativos.

3.3.1 Prosódia²⁵

Unidades entonacionais

Dependendo de sua complexidade, as contribuições dos falantes podem compreender uma ou mais unidades entonacionais.²⁶ Embora a segmentação de uma contribuição do falante em unidades entonacionais não seja completamente independente da sintaxe, a relação entre unidades entonacionais e unidades sintáticas não apresenta uma relação biunívoca. Isso se deve ao fato de que os participantes podem utilizar a segmentação de seus turnos para fins estilísticos, interacionais e semânticos (distinção entre foco e fundo, ver UHMANN, 1991; GUSSENHOVEN, 2004). Para a divisão da transcrição em segmentos, somente a prosódia, isto é, a segmentação em unidades entonacionais, é relevante.

Características e delimitações das unidades entonacionais

A unidade entonacional é percebida como uma unidade discreta por causa do seu contorno entonacional coesivo (“um trecho de fala proferido sob um mesmo contorno entonacional coerente”, DU BOIS et al., 1992, p. 17). A configuração final da unidade entonacional é de grande importância. O movimento entonacional final de um segmento é relevante, tanto para a percepção da delimitação da unidade entonacional, como para a interpretação da função interativa da unidade entonacional em questão. Uma unidade entonacional dispõe de pelo menos uma sílaba acentuada (o núcleo), isto é, uma sílaba que é foneticamente proeminente devido ao movimento entonacional, e/ou ao volume e à extensão,²⁷ e que crucialmente determina o significado da elocução. Em termos semânticos e pragmáticos, este acento tonal mais relevante será chamado *acento focal*.

Unidades entonacionais são separadas umas das outras por um número de sinais delimitadores opcionais. Esses podem ocorrer em combinação um com o outro, ou sozinhos. Os sinais individuais, por sua vez, podem ser mais ou menos salientes foneticamente. Na literatura (ver COUPER-KUHLEN, 1986; DU BOIS et al., 1992; CRUTTENDEN, 1997²⁸ etc., também SELTING, 1995), os seguintes sinais são mencionados:

²⁵ “Prosódia é um termo superordenado compreendendo os aspectos suprasegmentais da fala que resultam da interação dos parâmetros acústicos frequência fundamental (F_0), intensidade e duração no nível da sílaba ou em domínios mais amplos. Inclui fenômenos auditivos, como entonação, ou seja, a configuração entonacional da fala no decorrer do tempo, volume, extensão, pausa, bem como fenômenos de fala mais complexos, relacionados à velocidade/tempo e ao ritmo.” (SELTING, 1995, p. 1). Para um panorama do papel da prosódia na organização da fala-em-interação, ver também Couper-Kuhlen & Selting (1996), Ford & Couper-Kuhlen (2004), Couper-Kuhlen (2009), Selting (2010).

²⁶ Com relação à noção de ‘unidade entonacional’ ver Fox (2000) e Cruttenden (1997).

²⁷ Nota dos tradutores: A proeminência de movimento entonacional, volume e/ou extensão podem ocorrer alternada ou independentemente em uma mesma sílaba.

²⁸ Segunda edição.

- reajuste entonacional no início da unidade;
- maior velocidade da fala nas sílabas antecedendo a primeira sílaba acentuada (anacruse);
- prolongamento final ao fim da unidade;
- voz crepitante (*creaky voice*; glotalização) no final da unidade;
- movimento entonacional nas sílabas (não) acentuadas no final da unidade;
- pausas.

Os conceitos relevantes serão explicados detalhadamente a seguir.

Acento focal

Geralmente, unidades entonacionais contêm tanto sílabas acentuadas quanto não acentuadas, mas elas têm, pelo menos, uma sílaba acentuada. Sílabas acentuadas quase sempre se tornam proeminentes por meio de movimentos entonacionais (e possivelmente por meio de volume e duração). A percepção do acento é relacional, isto é, uma sílaba é acentuada ou não acentuada em comparação às sílabas à sua volta.

Os acentos em uma unidade entonacional são normalmente colocados em sílabas que são enfatizadas lexicalmente. A única exceção a isso é a acentuação contrastiva, como em: *eu disse meniNO e não meniNA* ou *é às TRÊS horas e não às TREze*. Todavia, nem toda sílaba lexicalmente enfatizada em uma unidade entonacional é acentuada.²⁹ Acento é um tipo de informação lexical, isto é, faz parte da entrada de uma palavra no vocabulário (*cor 'reto, 'hoje, ama 'nhã*). Em alguns casos, a ênfase até diferencia lexemas (*'dúvida – du 'vida; 'válido – va 'lido*).³⁰ Em algumas palavras, a ênfase pode variar.

Contudo, ainda que uma unidade entonacional possa ter vários acentos, normalmente apenas um deles possui relevância especial para o significado da elocução em questão. Este acento é chamado *acento focal*. O acento focal é o acento semântica e pragmaticamente mais relevante dentre as proeminências fonéticas reais na unidade entonacional. Ele indica e destaca o foco da elocução (ver abaixo). Na transcrição básica, sílabas com acento focal são indicadas por letras maiúsculas.³¹ A seguir, veremos alguns exemplos para uma notação correta das sílabas acentuadas:

1. com vogais tônicas orais: *CAsa, PEso, POSsa*

²⁹ Esta distinção entre ‘ênfase’ (potencial abstrato para proeminência) e ‘acento’ (proeminência foneticamente realizada) remete-nos a Bolinger (1964/1972); ver também Couper-Kuhlen (1986), Uhmman (1991, p. 21-22).

³⁰ Nota dos tradutores: Em português, nos casos em que ocorre a distinção de lexemas em função da ênfase, é comum que essa distinção seja marcada ortograficamente por sinal gráfico.

³¹ Esta convenção opõe-se àquela da AC e a de outros sistemas de transcrição. Porém, optamos por tal convenção uma vez que, letras maiúsculas (a) não precisam de comando de formato específico algum e, sendo assim, precisam de pouco tempo para serem escritas e (b) são geralmente conversíveis, isto é, elas também podem ser usadas em html.

2. com ditongos/tritongos: *PAI*, *linGUIça*, *UruGUAÍ*
3. com vogais tônicas nasais: *RIM*, *aTUM*

Quando um pico entonacional é mudado (foneticamente) para a sílaba anterior ou posterior à sílaba lexicalmente enfatizada, continua sendo a sílaba lexicalmente enfatizada que forma a sílaba acentuada (fonológica e semanticamente relevante): não percebemos estas mudanças como acentos mudados, mas só como contornos entonacionais variados.

Um acento excepcionalmente proeminente, sinalizado por movimento entonacional ou volume particularmente extensivo, é adicionalmente marcado por pontos de exclamação antes e depois da sílaba acentuada, como se vê na linha 56 do exemplo *Du ou Sie* (ver seção 7), por exemplo:

```

045   B3: e ela me pergun↑TOU,
→   046       (.) ah (-) perguntou se_ela po!DI!a me perguntAr:,
    047       ((ri)) se: <<sorrindo> eu era TURco.>

```

A posição do acento focal na unidade entonacional tem repercussões para a semântica da elocução. Isto se torna intuitivamente claro no seguinte conjunto de exemplos:

- a. eu vejo este texto na minha TEla
- b. eu vejo este TEXto na minha tela
- c. eu vejo ESTe texto na minha tela
- d. EU vejo este texto na minha tela

Todas estas acentuações produzem um foco diferente, isto é, elas sugerem significados e interpretações diferentes, que entendemos de forma intuitiva – mesmo sem saber as regras de colocação do acento. No conjunto de exemplos acima, (a) tem um foco abrangente, isto é, a sentença inteira está no foco. Em oposição, (b) a (d) têm um foco mais restrito ao sugerir, em cada caso, que haja uma alternativa ao elemento focalizado. Destarte, em (b), “(este) TEXTO” sugere uma alternativa como “não (esta) iMagem”, por exemplo. Em (d), o “EU” focalizado sugere “eu não voCÊ” ou “não Ele” como alternativas possíveis. Isso explica por que é essencial notar a acentuação na transcrição: caso contrário, mal-entendidos podem surgir com relação ao significado expresso pela elocução transcrita.³²

³² Em casos de dúvida, a sílaba com o acento focal pode ser identificada com ajuda do seguinte teste, que exclui possíveis alternativas de acentuações: produza, em voz alta, possíveis variantes de acentuações diferentes, compare-as umas com as outras, bem como com a gravação da elocução transcrita e decida qual delas combina melhor com a impressão auditiva no contexto sequencial relevante. Ao testar isso na linha 26 do trecho *Du ou Sie* (ver seção 7), a comparação mostrará que a posição apropriada do foco é:

sou mais velho que voCÊ

mais do que qualquer das possíveis alternativas, como, por exemplo:

SOU mais velho que você

Outros acentos dentro de uma unidade entonacional (que, por definição, não são acentos focais) não precisam ser indicados na transcrição básica. Se eles precisarem ser marcados, só o elemento vocálico da sílaba é que deve ser escrito em letras maiúsculas (e.g.: *este tExto*). Unidades entonacionais com vários acentos focais são raras, mas podem ocorrer. Um exemplo disso seria a resposta à pergunta: *quem inventou o quê?*

os chiNEses inventaram a acupunTura
mas os euroPEUS inventaram a cirurGIa

Observam-se, aqui, dois acentos focais em cada unidade entonacional.³³

Continuadores monossílabos são notados sem acento, isto é, com letras minúsculas. Se eles são mais altos do que a fala em volta, isso é indicado por meio de convenções descritas na seção 3.4., como, por exemplo, <<f> *hm*,>. Com continuadores dissílabos, a sílaba acentuada é notada em letras maiúsculas como, por exemplo, *hm_HM*.

Movimentos entonacionais no final das unidades entonacionais

Na transcrição básica, para cada unidade entonacional completa é estabelecido um símbolo indicando seu movimento entonacional final. Somente as unidades entonacionais incompletas não são notadas com símbolo algum.

Segundo as convenções de notação em AC, o movimento entonacional da unidade entonacional é notado com um dos símbolos a seguir:

- ? alto ascendente
- , ascendente
- nivelado³⁴
- ; descendente
- . baixo descendente

O símbolo para o movimento entonacional final é colocado logo após a última palavra da unidade entonacional; não há espaço entre as unidades entonacionais.

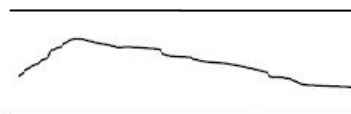
sou MAIS velho que você
sou mais VElho que você

As últimas transcrições sugerem focos para a elocução que não correspondem à impressão auditiva ou ao contexto da narrativa e, por conseguinte, parecem menos plausíveis.

³³ Ver Couper-Kuhlen (1986, p. 47-48), Ladd (2008, p. 7-8), também Uhmman (1991, p. 221).

³⁴ O símbolo ‘–’ não é um sinal de menos ou um hífen, mas sim um travessão (U+2013). É necessário marcar o nível tonal, porque os símbolos de movimento tonal final indicam concomitantemente o fim da unidade entonacional. Portanto, nenhuma unidade entonacional completa deveria ficar sem tal símbolo, ainda que o tom não mude, como no caso do tom nivelado. Em tais casos, o fim da unidade entonacional é geralmente marcado com outros parâmetros como intensidade e duração.

O último movimento entonacional pode se estender ao longo de um número variado de sílabas. Na maioria dos casos, o movimento entonacional começa no acento focal da sílaba e continua até o fim da unidade. Nessas ocorrências, o tom gradualmente desce, ascende ou permanece nivelado do acento focal em diante. Este é o caso do movimento entonacional descendente no exemplo a seguir, retirado da narrativa do *Du ou Sie* (ver seção 7):



29 B2: (...) NÃO ia reparar.

No caso de um contorno descendente-ascendente, ou seja, quando o tom cai do acento focal em diante e quando ascende novamente, apenas o movimento entonacional final é notado:



46 B3 : (...) poDIA me perguntar,

No exemplo acima, o tom, a princípio, cai de um pico silábico *DI-* sobre a sílaba seguinte *-a* antes de inverter a direção na palavra *perguntar* e ascender ao grau médio na parte final do segmento. Em outras palavras, ainda que, no geral, o movimento entonacional final inteiro, a partir do acento focal, seja descendente e, então, ascendente, somente o tom ascendente é notado, pois a transcrição básica registra apenas o movimento entonacional final ao fim do segmento.³⁵

É importante notar que os símbolos usados aqui não são sinais de pontuação. Como consequência disso, eles não indicam a função de orações ou elocuições como, por exemplo, afirmações ou perguntas; eles se referem apenas ao movimento entonacional final da unidade entonacional.

Unidades interrompidas que não acabam com fechamento glotal, mas simplesmente com *trail-off*, não são notadas com símbolo de entonação final algum:

x: eu pensei eu posso

³⁵ Para notar o movimento tonal mais precisamente, as convenções de notação da transcrição refinada podem ser adicionadas (ver seção 3.4.). Isso faz com que seja possível representar os movimentos descendentes e ascendentes em sua distribuição sobre as sílabas acentuadas:

po´DIA me perguntar,

Aqui, é a notação sem o símbolo para a entonação final que indica que a unidade entonacional está incompleta.³⁶

O símbolo ‘-’ representa que a entonação final permanece em nível no meio da extensão de som do falante.³⁷

Note que elementos de começo de turno e *tags* realizam-se, às vezes, como unidades entonacionais separadas e independentes prosodicamente, com seus próprios movimentos entonacionais finais. Em casos de *tags* com movimentos entonacionais separados, tanto o movimento entonacional ao fim da unidade entonacional precedente como o movimento entonacional ao fim da própria *tag* são notados. A *Tag* e a unidade entonacional são geralmente conectadas por engatamento, o qual é notado com ‘=’ (ver 3.3.2.):

01 A: aQUI é onde a transcrição começa;=né?

Ao mesmo tempo é possível realizar a *tag* como uma unidade entonacional independente:

01 A: aQUI é onde a transcrição começa.

02 né?

Ou ambos os elementos podem ser completamente integrados em uma só unidade entonacional:

01 A: aQUI é onde a transcrição começa né?

Elementos iniciadores e marcadores discursivos prefaciadores de turno, tais como *então*, *sim*, *não*, *bem*, *okay*, também podem ser notados como (a) itens não acentuados pertencentes à unidade entonacional seguinte, (b) unidades entonacionais prosodicamente dependentes, ou (c) unidades entonacionais prosodicamente independentes segundo sua realização. Compare o marcador de prefácio de turno *então* e a *tag né* nas três versões seguintes, com limites gradualmente mais marcados entre *então* e o sintagma seguinte, bem como entre *né* e o sintagma precedente.³⁸

³⁶ Se necessário para objetivos específicos de pesquisa, o símbolo especial ‘\’ (U+005C) pode ser inserido quando unidades entonacionais são quebradas sem fechamento glotal.

³⁷ Para uma notação mais detalhada do nível de entonação ao fim de uma unidade entonacional, veja as convenções para transcrição refinada.

³⁸ Em geral, os transcritores podem treinar sua percepção por meio de repetições em voz alta das elocuições para si mesmos, auto-observação e, em particular, ao comparar suas próprias produções com as elocuições gravadas. Ademais, alguns testes simples podem ajudar a identificar o movimento tonal para a transcrição: para esclarecer se o tom ao final da unidade entonacional no segmento 40 de nosso exemplo *Du ou Sie* é descendente, como notado,

029 B2: NÃO ia reparar.

- a. 01 A: então aQUI é onde a transcrição começa né?
- b. 01 A: enTÃO;=aQUI é onde a transcrição começa;=né?
- c. 01 A: enTÃO;
 02 aQUI é onde a transcrição começa;
 03 né?

3.3.2. Outras convenções para transcrição básica.

Falas engatadas (Latching das unidades entonacionais)

= Fala engatada, ou seja, mudança imediata para uma nova unidade entonacional (pelo mesmo falante ou por outro) sem a micropausa comum.

O símbolo de engatamento é colocado ao final da unidade entonacional precedente e no início da próxima. Por exemplo:

01 A: eu tamBÉM;=
 02 B: =eu tamBÉM;

Nas unidades entonacionais, as palavras geralmente são produzidas diretamente uma após a outra. Por isso, a fala engatada, nesse caso, é a regra, e conseqüentemente não é notada.

Alongamento

: :: ::: deve ser usado de acordo com a duração do alongamento, por exemplo, *então:* ou *é::*, etc.

Os dois pontos são usados como seguem:

: alongamento de aproximadamente 0,2 até 0,5 segundos.
 :: alongamento de aproximadamente 0,5 até 0,8 segundos.
 ::: alongamento de aproximadamente 0,8 até 1 segundo.

produzir a elocução em voz alta com tom ascendente, ao invés de um tom descendente, e compará-lo com os registros originais pode rapidamente resolver a questão. Alunos avançados com conhecimentos de fonética acústica podem também querer usar um programa de fonética acústica para a análise de fala tal como o PRAAT (URL: <<http://www.praat.org>>).

A percepção da duração dos sons ou sílabas alongadas depende, por sua vez, da acentuação, tempo e ritmo.³⁹

Fechamento glotal

? ruptura com fechamento glotal ou início de sílaba com fechamento glotal (U+0294)⁴⁰

Não é notado o uso de fechamento glotal que esteja de acordo com as regras de pronúncia de uma língua (ver seção 3.2.1.). (Ver continuadores na seção 3.2.3.)

Comentários interpretativos

<<irritado> > } comentários interpretativos com a indicação de seu escopo
<<surpreso> > }

Fenômenos auditivos que o transcritor não consegue representar em termos da forma, mas que são presumidamente relevantes para a interpretação, podem ser notados inicialmente por meio de comentários interpretativos. Nesses casos, comentários como ‘<<surpreso> >, <<chateado> >, <<hesitante> >’ etc., são localizados no início da fala relevante (ou em parte da elocução). O parêntese angular exterior deve ser colocado no lugar, em que o escopo com a característica descrita pelo comentário finaliza. O escopo de um comentário pode variar de segmentos fonéticos a sílabas, palavras, frases, unidades entonacionais ou até turnos inteiros. Se um comentário for válido para mais de um turno, deve ser repetido no início do novo turno:

01 A: você é <<hesitante> talvez um policial
02 será que eu te conheço da delegacia>
03 B: desculpe
04 A: <<hesitante> você é da polícia>

Comentários interpretativos podem ser completados ou substituídos em um estágio posterior por descrições mais formais.

³⁹ O encurtamento não é notado por um símbolo diferente. Se necessário, uma realização mais curta da palavra inglesa /ru:f/ ‘roof’ (teto, N. T.), como /rUf/ por exemplo, pode ser indicada por *ruf*.

⁴⁰ Note-se que o símbolo fonético do fechamento glotal é sobrescrito. Se este símbolo não estiver disponível ou não for confiável, o símbolo * (U+002A) pode ser usado em substituição. Se um símbolo para o rompimento sem fechamento glotal for necessário para fins específicos, \ (U+005C) pode ser usado.

3.4. A transcrição refinada

Dependendo da necessidade de cada um, a transcrição básica pode ser expandida pela adição de outros parâmetros para análise. Esta seção apresenta outras convenções que são relevantes para a transcrição de prosódia.

Além de uma indicação mais precisa do posicionamento e da ênfase dos acentos, a transcrição prosódica refinada inclui o registro de movimentos entonacionais nas sílabas enfatizadas e depois delas, de pulos entonacionais perceptíveis no início ou ao longo das unidades entonacionais, de mudanças no registro entonacional, assim como no volume e na velocidade. São mantidos os parâmetros prosódicos da transcrição básica.

*Colocação de acento e ênfase:
acento focal e outros acentos (acentos secundários)*

Sílabas que trazem um acento focal (ou acento primário) são notadas com letras maiúsculas – como na transcrição básica (ver seção 3.3.1). Na transcrição refinada, quaisquer outros acentos adicionais também são notados. Referimo-nos a eles como acentos secundários. Em comparação com os acentos focais, os acentos secundários são semântica e pragmaticamente menos relevantes, e frequentemente também são menos proeminentes. Como os outros acentos, eles também se manifestam em sílabas, e, em palavras multissilábicas, geralmente caem na sílaba lexicalmente acentuada (ver 3.3.1). Nas sílabas que apresentam um acento secundário, notamos com letra maiúscula apenas o grafema núcleo da vogal principal (e.g.: sozInho, hOje, sazonAl, peÃO, cAuda, fEira).⁴¹

Sílaba	acento primário
sílaba	acento secundário

O exemplo a seguir ilustra a notação de acento focal (AF) e de acento secundário (AS):

susAna pediu mantEiga e uma fatia de PÃO
AS AS AF

*Pulos entonacionais perceptíveis no início ou ao longo de uma unidade entonacional
(em comparação com a unidade entonacional ou a sílaba anterior)*

↑	pequeno pulo entonacional para cima (U+2191)
↓	pequeno pulo entonacional para baixo (U+2193)

⁴¹ Essa diferenciação ortográfica não é possível em palavras que contêm a sílaba tônica com apenas um grafema, como *Ovo*, *Ida*, *Olho* etc.

↑↑ grande pulo entonacional para cima
↓↓ grande pulo entonacional para baixo

Pulos entonacionais perceptíveis são mudanças claras e repentinas na frequência do som no início de uma sílaba, em comparação com a frequência do som nas sílabas anteriores, acentuadas ou não acentuadas. Eles são notados no momento em que ocorrem, isto é, tanto no início da unidade entonacional, quando ela começa, com uma frequência perceptivelmente mais alta do que as anteriores, quanto ao longo da unidade entonacional, quando, por exemplo, os acentos forem realizados com pulos para picos ou vales na curva entonacional, os quais se destacam como altos ou baixos demais em relação aos acentos anteriores. Tais pulos entonacionais são fenômenos locais, se contrastados às mudanças na frequência do som discutidas abaixo. O símbolo de notação desses pulos deve ser colocado à frente da sílaba, à qual se refere.

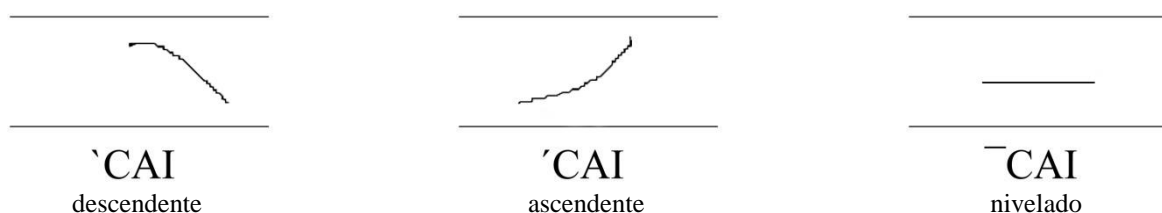
Mudanças na frequência do som

<<l> > frequência mais baixa
<<h> > frequência mais alta

Esses símbolos são usados quando o falante muda para uma frequência de som que difere da anterior ou da frequência normalmente utilizada por este falante – em uma ou mais unidades entonacionais, ou em partes de uma unidade entonacional, tal como em inserções parentéticas. O escopo da mudança na frequência do som pode variar, mas normalmente se aplica a trechos maiores em um turno (para mais informações sobre o uso de parênteses angulares, ver seção 3.2.4. e paralelos).

Acentos tonais

A notação comum de acentos tonais indica a curva entonacional na sílaba enfatizada e após esta sílaba. Normalmente, deve ter uma das seguintes formas:



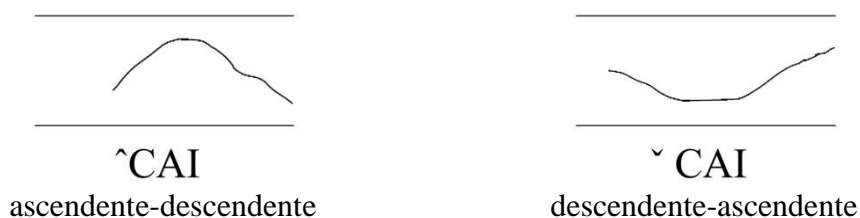
Quando picos e vales da curva entonacional são produzidos no meio da sílaba enfatizada, por exemplo, no segmento fonético /aI/ (“ai”) deste exemplo, a notação do acento indica a curva entonacional que começa aqui e continua ao longo do restante da sílaba, e frequentemente se espalha também pela palavra e pelas sílabas não acentuadas que seguem. Portanto, os acentos

tonais da palavra “cai” mostrados acima são, respectivamente, descendente, ascendente e nivelado. O símbolo relevante é colocado à frente da sílaba acentuada:⁴²

en`TÃO	descendente (U+0060)
en^TÃO	ascendente (U+00B4)
en-TÃO ⁴³	nivelado (U+00AF)

Uma mudança de parâmetro em comparação à sílaba anterior, tal como, por exemplo, um pulo para cima ou para baixo em direção à sílaba enfatizada, ou um deslizamento para cima em direção ao pico ou para baixo em direção ao vale do acento tonal é indicado por setas (ver abaixo).

Além disso, há casos em que são produzidas curvas descendentes-ascendentes ou ascendentes-descendentes em uma simples sílaba enfatizada, o que frequentemente a torna mais longa. Tais casos estão esquematizados nas figuras seguintes:



Tais acentos tonais descendente-ascendente e ascendente-descendente são representados da seguinte forma:

en^TÃO⁴⁴ ascendente-descendente (U+02C6)

en^vTÃO⁴⁵ descendente-ascendente (U+02C7)

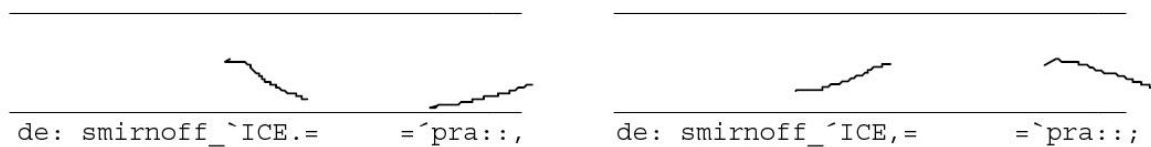
Ainda há outros casos em que a primeira parte de um acento tonal ascendente-descendente ou descendente-ascendente é realizada na sílaba enfatizada, mas a segunda parte ocorre apenas depois, em uma sílaba não enfatizada:

⁴² Os símbolos utilizados aqui são os acentos gráficos agudo e grave do teclado; não são apóstrofes ou aspas simples.

⁴³ Alternativamente: en-TÃO (com U+2013).

⁴⁴ Alternativamente: en^TÃO (uma combinação dos símbolos para acento ascendente e descendente, i.e. U+00B4 e U+0060).

⁴⁵ Alternativamente: en^vTÃO (uma combinação dos símbolos para acento descendente e ascendente, i.e. U+0060 e U+00B4).



Nesses casos, os movimentos entonacionais são notados no lugar em que ocorrem, isto é, a primeira parte da curva entonacional na sílaba enfatizada, e a segunda parte, na sílaba não enfatizada relevante.

Sendo assim, combinações da notação do tipo de acento e do movimento entonacional vão aparecer das seguintes maneiras, por exemplo:

Acento focal:
descen`DENte
ascen´DENte
nive¯LAdo

Acento secundário:
descen`dEnte
ascen´dEnte
nive¯lAdo

Pulos para cima ou para baixo no início da sílaba enfatizada, assim como diferenças entre as sílabas anteriores não enfatizadas e o pico ou o vale da sílaba enfatizada podem ser representados da seguinte forma:⁴⁶

↑`	pequeno pulo entonacional para cima em direção ao pico da sílaba enfatizada (U+2191 e U+0060)
↓´	pequeno pulo entonacional para baixo em direção ao vale da sílaba enfatizada (U+2193 e U+00B4)
en↑↑`TÃO ou en↓↓´TÃO	pulos entonacionais perceptíveis em direção ao pico ou ao vale da sílaba enfatizada
en↑¯TÃO ou en↓¯TÃO	pulos entonacionais para acentos perceptivelmente mais altos ou mais baixos com curva nivelada

Em combinação com a notação do movimento no fim da unidade entonacional (ver 3.3.1.), as convenções sugeridas acima captam o quão alto ou baixo foi o último movimento da unidade entonacional: um movimento descendente final pode cair para o meio (notado com ‘;’)

⁴⁶ Se necessário, o símbolo para entonação nivelada no fim de uma unidade entonacional (¯ = U+2013, ver transcrição básica, 3.3.) pode ser inserido como subscripto para representar uma entonação nivelada em uma frequência mais baixa (¯_ = U+005F) ou como sobrescrito (¯^ = U+00AF) para representar uma entonação nivelada em uma frequência mais alta.

ou para baixo (‘.’), já um movimento ascendente final pode se elevar ao meio (‘,’) ou ao alto (‘?’). Observe:

001 B1: teve_uma vez que a gente tAva no quiosque comprando cer´VEja,
042 B3: uma menina (.) °h uma menina achou que_eu era ´↑TURco?

No primeiro exemplo, a curva entonacional se eleva ao meio nas sílabas finais enfatizadas, já no segundo, eleva-se ao alto no final do acento secundário.

Mudanças de volume e velocidade

A notação do volume e da velocidade só se torna relevante, quando ocorre uma mudança em relação ao volume e à velocidade estabelecidos para os trechos de fala anteriores. Se um falante fala habitualmente alto ou rápido, isso deve ser indicado no cabeçalho da transcrição. Neste caso, o hábito do falante é o ponto de referência para a transcrição. Mudanças no volume e na velocidade são notadas com tanto detalhamento quanto possível, de acordo com os seguintes parâmetros:

<<f>	>	forte, alto ⁴⁷
<<ff>	>	fortissimo, muito alto
<<p>	>	piano, baixo
<<pp>	>	pianissimo, muito baixo
<<all>	>	allegro, rápido
<<len>	>	lento, devagar
<<cresc>	>	crescendo, aumentando o volume
<<dim>	>	diminuendo, diminuindo o volume
<<acc>	>	acelerando, aumentando a velocidade
<<rall>	>	rallentando, diminuindo a velocidade

A indicação de uma mudança no volume ou na velocidade se refere a um trecho de fala inserido entre os parênteses: o descritor é posicionado (também em parênteses angulares) antes da sílaba, em que a mudança do parâmetro pode ser observada; o parêntese mais externo é colocado no momento em que termina o escopo para tal parâmetro.⁴⁸

⁴⁷ O GAT não pode utilizar a convenção da AC para notar o volume alto, letras maiúsculas, já que estas são utilizadas para indicar a ênfase. O fenômeno prosódico da ênfase ocorre não apenas devido ao volume da voz: para identificar ênfases, é necessário levar em conta a duração, e em especial a curva entonacional, além do volume.

⁴⁸ Quando forem empregados parâmetros com diferentes escopos, os parênteses angulares podem ser utilizados com índices, tais como <<f, h> f > h>.

Mudanças na qualidade vocal e no modo de articulação

- <<crepitante> > glotalizado, “voz rangida”
<<sussurrante> > exemplos de mudança de qualidade vocal como indicado, por exemplo, para fonação sussurrante.

Tais parâmetros podem ser usados também para trechos do turno de extensão variável. Assim como os parâmetros de volume e tempo descritos acima, estes são delimitados por parênteses angulares. Outros descritores comuns para mudanças de qualidade vocal incluem, por exemplo, ofegante, dura, relaxada, tensa, em falsete, com lábios arredondados, com lábios distensos, nasalizada, desnasalizada etc. (ver LAVER, 1980).

Outros fenômenos fonéticos

- ((clique))
((estalo))

((clique labial))
((clique alveolar))
- } ações e eventos fonológicos, para- e não linguísticos (os descritores são utilizados de modo tão detalhado quanto possível)

Sinais limítrofes

A fim de notar os sinais limítrofes nas unidades entonacionais conforme descrito acima, os seguintes símbolos podem ser usados:

- = fala engatada
↑↓ reajuste entonacional no início de unidades entonacionais
<<all> > aceleração da fala nas primeiras sílabas da unidade (anacruse)
: prolongamento terminal no fim da unidade
<<crepitante>> voz crepitante no fim da unidade
. ; - , ? movimentos entonacionais finais nas últimas sílabas acentuadas (e não acentuadas)
(.) etc. micropausa ou outro tipo de pausa após uma unidade ou entre unidades
°h h° inspiração e expiração após uma unidade ou entre unidades

4. Sugestões para a notação de fenômenos complexos

Nesta seção, apresentamos sugestões para estender a transcrição refinada, particularmente em termos de entonação e ritmo, para problemas de pesquisa específicos.

4.1. Representação autosegmental da entonação

Como uma alternativa para o registro auditivo/impressionista da entonação, tal como apresentado na seção 3, o GAT 2 permite a integração de uma representação autosegmental dos contornos entonacionais em transcrições básicas e refinadas para fins específicos de pesquisa.⁴⁹ Tais representações normalmente requerem uma análise *a priori* da entonação e da criação de um inventário de tons e sequências entonacionais para cada variedade linguística pesquisada. Tons e sequências entonacionais específicos desse inventário são atribuídos a segmentos de transcrição individuais. Via de regra, a representação autosegmental implica uma abordagem fonológica, isto é, os contornos entonacionais observados são entendidos como instâncias de padrões abstratos no inventário de tons de uma determinada língua. No entanto, a notação autosegmental também pode ser usada para as variedades, cujo inventário entonacional ainda não é conhecido. Logo, os símbolos usados pelo ToBI representam, antes de tudo, uma notação fonética rigorosamente abstrata; eles são apenas potencialmente fonológicos. Para uma notação exclusivamente fonética, veja Grabe et al. (1998) e também Gilles (2005).

A seguir, a integração da representação autosegmental no GAT é ilustrada por meio de convenções de notação do ToBI para a entonação das variedades do Português Brasileiro e Europeu (FROTA et al., 2015a). Está além do presente trabalho introduzir o ToBI em detalhe, tampouco refletir sobre as tendências atuais sobre sua discussão (cf. LADD, 2008). Em vez disso, o objetivo é apontar para a possibilidade de integração de (alguns) dos seus conceitos relevantes.

Uma notação ‘adequada’ do ToBI inclui várias camadas (cf. GRICE, 2006; LADD, 2008). Nós nos concentraremos aqui, para exemplificação, na camada entonacional. Contornos entonacionais são entendidos como sequências de eventos entonacionais locais (alvos) e transições entre estes. Os alvos entonacionais são especificados em termos de tons altos (H) e tons baixos (L). As transições geralmente evoluem a partir da interpolação linear dos alvos entonacionais.

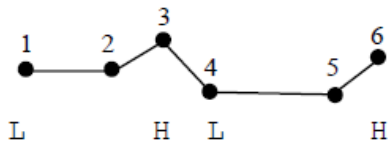
Um contorno entonacional é uma sequência de quaisquer tons H e L. Esquema (1) mostra uma dada sequência entonacional; os pontos pretos representam alvos fonéticos de tons individuais.



Esquema 1

⁴⁹ Para uma introdução à representação autosegmental da entonação, ver Ladd (2008), Grussenhoven (2004) e Beckman & Ayers Elam (1997).

Alvos vizinhos que são ambos baixos ou altos podem ser associados, respectivamente, com um único tom baixo, ou alto (dispersão entonacional). Desse modo, em (2), o primeiro tom L especifica o primeiro e o segundo alvos, e o segundo L, o quarto e o quinto alvos.



Esquema 2

Existem quatro tipos de tom: tons acentuais, tons condutor/sucessor, tons frasais e tons limítrofes. Os tons acentuais (H^* , L^*) constituem o núcleo dos acentos entonacionais. Estes estão associados a sílabas proeminentes, sendo a associação sinalizada por um asterisco. Acentos monotonais consistem de um único tom acentual. Os tons condutor e sucessor (H , L) ocorrem somente em combinação com tons acentuais. Juntamente com um tom acentual, eles formam acentos bitonais. Dependendo se eles precedem ou sucedem o tom acentual a que pertencem, eles são chamados de condutor ou sucessor, respectivamente. Os tons frasais e limítrofes são tons marginais. Os tons frasais (H ou H^- , L ou L^-) estão associados à margem de um constituinte prosódico, chamado unidade intermediária,⁵⁰ os tons limítrofes são associados às margens de unidades entonacionais. Os tons limítrofes que indicam um início marcadamente alto de uma unidade entonacional são notados por meio de $\%H$, aqueles que estão no fim de uma unidade entonacional são notados ou com $L\%$, ou com $H\%$.

Dependendo da versão específica do ToBI⁵¹, há quatro tipos de acentos tonais, os quais podem ocorrer majoritariamente como acentos nucleares em Português Brasileiro: um acento descendente ($H+L^*$), um acento côncavo (L^*+H), um acento ‘pico’ ascendente ($L+H^*$) e um acento baixo (L^*). No ToBI, o acento nuclear é, por definição, o último acento tonal de uma unidade entonacional. Quaisquer acentos precedentes são chamados acentos pré-nucleares. Combinações especificadas de um dos quatro acentos tonais $H+L^*$, L^*+H , $L+H^*$ e L^* em posição nuclear juntamente com um dos tons frasais H , L e um dos tons limítrofes finais $L\%$, $H\%$ resultam em sequências entonacionais, que correspondem aproximadamente aos acentos descendente, ascendente, nivelado, descendente-ascendente e ascendente-descendente, descritos para o GAT nas seções 3.3.1 e 3.4 acima.⁵²

Encontra-se, aqui, uma tabela de correspondências para o inventário de acentos nucleares (marcas de acento entonacional que são finais) em Português Brasileiro (FROTA et al, 2015a):

⁵⁰ O termo ‘unidade intermediária’ remete ao termo em inglês *intermediate phrase*. A tradução de *phrase* para ‘unidade’ se deve, no caso, a uma questão de coerência textual, sobretudo no que se refere à tradução de *intonational phrase* para ‘unidade entonacional’. Para uma reflexão mais completa acerca dessa problemática terminológica, ver Barth-Weingarten (2011).

⁵¹ Aqui nós seguimos P_ToBI (Portuguese ToBI, ver FROTA et al., 2015b).

⁵² Para mais exemplos em Português, o leitor é referenciado a Frota et al. (2015a).

Acento tonal	Acento frasal ⁵³	Acento limítrofes
H+L*	?	L%
H+L*	?	H%
H+L*	?	HL%
L*+H	L ₋	L%
L*+H	?	H%
L*+H	?	HL%
L+H*	?	L%
L+H*	?	!H%
L*	?	L%
L*	?	HL%

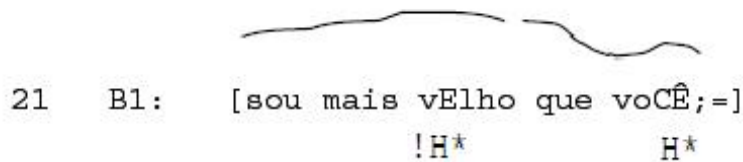
Em casos não marcados, acentos nucleares correspondem a acentos focais no GAT, e acentos pré-nucleares, a acentos secundários no GAT (ver seção 3). A esse respeito, sílabas com acentos nucleares são notadas com letras maiúsculas na linha de notação ortográfica. Em sílabas com acentos pré-nucleares, somente o núcleo silábico é capitalizado; em sílabas com acentos pré-nucleares contendo ditongos ou várias vogais, somente a primeira vogal é capitalizada. O contorno entonacional ToBI é notado em uma linha separada, abaixo da linha relevante da transcrição ortográfica do GAT. Ambas as linhas formam um segmento de transcrição. Por isso, à linha que indica a entonação não é designado um número de segmentação separado.

Em termos de fonte e tamanho, a linha de entonação assemelha-se à linha de notação ortográfica. Os tons marcados com asterisco são notacionalmente alinhados tanto quanto possível com as vogais das sílabas acentuadas. Os tons limítrofes são alocados de tal modo que a letra que indica o tom (não o sinal de porcentagem) é alinhada com o começo ou o final da linha de notação verbal relevante, respectivamente. Para uma ilustração, veja o seguinte excerto de *Du ou Sie*:

021 B1: <<sorrindo> sou mais vElho que voCÊ;>=
!H* H*

Para propósitos de ilustração, segmentos de transcrição que consistem de uma linha de notação ortográfica e uma linha de notação entonacional podem também ser complementados por curvas entonacionais impressionistas, como indicado abaixo. Estas últimas são alocadas acima da respectiva linha de transcrição verbal. Tais curvas entonacionais não são parte do segmento de transcrição.

⁵³ Em Frota et al. (2015a), há somente referência a esse tipo de acento frasal, o qual indica tendencialmente, segundo os autores, *Narrow Focus Statement*.



A curva entonacional serve para facilitar o entendimento da notação entonacional, não para indicar algum detalhe fonético adicional. Para tanto, extrações de F0 devem ser utilizadas; estas, contudo, não são integradas na transcrição.⁵⁴

4.2. Notação de ritmo

Ritmo pode emergir quando sílabas acentuadas são pronunciadas isocronamente, isto é, quando seguem umas às outras em intervalos de tempo perceptualmente iguais, ou quando duas unidades entonacionais imediatamente adjacentes e igualmente longas exibem lugares de acentuação recorrentes. Os diferentes princípios que resultam em estrutura rítmica necessitam de diferentes notações.⁵⁵

Uma unidade de ritmo básica é o *pé* (ver NESPOR & VOGEL, 2007). Ela inclui uma sílaba acentuada, percebida como uma batida rítmica, e qualquer sílaba não acentuada seguinte até a próxima sílaba acentuada (esta é parte do próximo pé). Para a representação do pé rítmico produzido isocronamente, nós recomendamos a notação desenvolvida por Couper-Kuhlen (1993, ver também AUER et al., 1999). Usando esse sistema, as sílabas acentuadas na unidade entonacional *do /tIpo uma /COIsa assim;/=sAbe*. (linha 92 no exemplo GAT em português *Du ou Sie*; ver seção 7), que são percebidas como isócronas, seriam notadas da seguinte maneira:

063 B3: do /tIpo uma /
/COIsa assim;= /
/=sAbe. /

Cada pé é notado em uma linha separada. Os cortes à esquerda e à direita, os quais estão alinhados uns aos outros, indicam a duração perceptualmente igual de cada pé. A efetiva

⁵⁴ Deve-se considerar, em conclusão, que o ToBI, bem como o GAT, precisa ser adaptado à língua ou variedade linguística particular sob investigação.

⁵⁵ Dado o estado atual de pesquisa sobre o ritmo da fala, somente o princípio rítmico de isocronia acentual será apresentado aqui. Uma discussão do princípio rítmico de distribuição acentual recorrente (em Alemão com influências de Turco) bem como uma representação proposta baseada em grades métricas (ver SELKIRK, 1984) podem ser encontradas em Kern (2008). Outros estudos sobre a ocorrência e função desse princípio em outras variedades linguísticas e estilos de fala precisam ainda ser feitos.

ou turno logo abaixo do segmento de transcrição verbal na língua original. Esta linha de tradução deve estar na fonte *Times New Roman*. Caso o estudo considere aspectos tipológicos relevantes, pode ser desejável a inserção de uma glosa interlinear entre a transcrição verbal na língua original e a tradução livre. A glosa interlinear normalmente indica a estrutura morfológica do original (ver CROFT, 2003; LEIPZIGGLOSSING RULES, 2008).

Isso será ilustrado a partir de um exemplo do finlandês (SORJONEN, 2001, p. 413). A primeira linha traz a transcrição original, a segunda traz a glosa interlinear e a terceira, a tradução livre:

- 5 E: Meinaa-k se nyt jää-hä kotti-i sitte.
Intend-Q it now stay-INF home-ILL then
 Is she gonna stay at home then.
- 6 M: Joo::, Joo::,
PTCL PTCL

Na glosa interlinear, cada morfema da elocução original (L1; aqui, finlandês) é representado por um morfema próprio em L2 (aqui, inglês). Os morfemas lexicais da L1 são traduzidos em morfemas lexicais da L2 e apresentados em itálico na linha abaixo (e.g. na linha 5, *meinaa* = ‘intend’; *jää* = ‘stay’). Os morfemas gramaticais da L1 são glosados de acordo com um conjunto de categorias gramaticais determinadas metalinguisticamente (ver tabela abaixo) representadas por abreviações em letras maiúsculas (e.g. *k* = Q (partícula interrogativa), *i* = ILL (ilativo ‘dentro’)) etc. Os limites dos morfemas são indicados pelo sinal de menos (-), e limites de palavras, por espaços em branco (). Partículas e outras palavras de difícil tradução, tal como, por exemplo, *joo* na linha 6, podem permanecer sem tradução, principalmente quando são o objeto de estudo.

A seguir há uma lista de categorias gramaticais comuns junto às suas respectivas abreviações (CROFT, 2003). Pode ser necessário suplementá-la, a depender da língua em questão:

<i>Categoria gramatical</i>	<i>Abreviação</i>
Definido	DEF
Indefinido	INDEF
Masculino	M
Feminino	F
Neutro	N
Nominativo	NOM

Genitivo	GEN
Dativo	DAT
Acusativo	ACC
Singular	SG
Plural	PL
Diminutivo	DIM
Forma familiar (T)	FAM
Objeto direto	DO
Objeto indireto	IO
Presente	PRS
Pretérito	PST
Primeira pessoa	1
Segunda pessoa	2
Terceira pessoa	3
Verbo auxiliar	AUX
Partícula	PTCL
Infinitivo	INF
Particípio	PART
Subjuntivo	SUBJ
Voz passiva	PASS
Negação	NEG
Complemento	COMP
Quotativo	QUOT

6. Sumário das mais importantes convenções de transcrição do GAT2

Transcrição mínima

Sequência estrutural

[] sobreposição e fala simultânea
[]

Inspirações e expirações

°h/h° ins-/expiração de aprox. 0,2-0,5 seg de duração
°hh/hh° ins-/expiração de aprox. 0,5-0,8 seg de duração
°hhh/hhh° ins-/expiração de aprox. 0,8-1,0 seg de duração

Pausas

(.) micro pausa estimada em até 0,2 seg de duração aprox.
(-) pausa curta estimada em aprox. 0,2 – 0,5 seg de duração
(--) pausa intermediária estimada em aprox. 0,5 – 0,8 seg de duração
(---) pausa longa estimada em aprox. 0,8 – 1,0 seg de duração
(0.5)/(2.0) pausa mensurada em aprox. 0,5/2,0 seg de duração (até o décimo de segundo)

Outras convenções segmentais

e_ah cliticizações dentro de unidades
eh, ahm, uhm etc. marcadores de hesitação, assim chamados de “pausas preenchidas”

Risos e choros

haha }
hehe }
hihi } risada silábica

((ri)) }
((chora)) } descrição de risada ou choro

<<rindo> > partículas de riso acompanhando a fala com indicação de escopo
<<:-> então> “smile voice”

Continuadores

hm, mm, sim, é	partículas monossilábicas
hm_hm, aham	partículas bissilábicas
ʔhmʔhm	com fechamento glotal, frequentemente para negação

Outras convenções

((tosse))	ações e eventos vocais não verbais
<<tossindo> >	descrição com indicação de escopo
()	trecho incompreensível
(xxx), (xxx xxx)	uma ou duas sílabas incompreensíveis
(posso)	termo presumido
(posso/passo a falar)	possíveis alternativas
((incompreensível, aprox. 3 seg))	trecho incompreensível com indicação de duração
((...))	omissão na transcrição
→	refere-se a uma linha de transcrição relevante na argumentação

Transcrição básica

Sequência estrutural

= continuação rápida e imediata com um novo turno ou segmento (*latching*)

Outras convenções de segmentação

:	alongamento, de aprox. 0,2-0,5 seg.
::	alongamento, de aprox. 0,5-0,8 seg.
:::	alongamento, de aprox. 0,8-1,0 seg.
ʔ	ruptura (<i>cut-off</i>) por fechamento glotal

Acentuação

Sílaba	acento focal
!SÍlaba	acento focal extraforte

Movimentos entonacionais no final das unidades entonacionais

ʔ	alto ascendente
,	ascendente
-	nivelado

; descendente
. baixo descendente

Outras convenções

<<surpreso> > comentário interpretativo com indicação de escopo

Transcrição refinada

Acentuação

Sílaba acento focal
sílaba acento secundário
!SÍ!laba acento extra focal

Pulos entonacionais

↑ pequeno pulo entonacional para cima
↓ pequeno pulo entonacional para baixo
↑↑ grande pulo entonacional para cima
↓↓ grande pulo entonacional para baixo

Mudanças na frequência do som

<<1> > frequência mais baixa
<<h> > frequência mais alta

Notação intralinear de movimentos entonacionais

en`TÃO descendente
en´TÃO ascendente
en⁻TÃO nivelado
en[^]TÃO ascendente-descendente
en[~]TÃO descendente-ascendente

↑` pequeno pulo entonacional para cima em direção ao pico da sílaba enfatizada
↓´ pequeno pulo entonacional para baixo em direção ao vale da sílaba enfatizada

en_↑ˉTÃO ou en_↓ˉTÃO pulos entonacionais para acentos perceptivelmente mais altos ou mais baixos com curva nivelada

en_↑↑`TÃO ou ↓↓´TÃO pulos entonacionais perceptíveis em direção ao pico ou em direção ao vale da sílaba enfatizada

Mudanças de volume e velocidade, com escopo

<<f>	>	forte, alto
<<ff>	>	fortissimo, muito alto
<<p>	>	piano, baixo
<<pp>	>	pianissimo, muito baixo
<<all>	>	allegro, rápido
<<len>	>	lento, devagar
<<cresc>	>	crescendo, aumentando o volume
<<dim>	>	diminuendo, diminuindo o volume
<<acc>	>	acelerando, aumentando a velocidade
<<rall>	>	rallentando, diminuindo a velocidade

Mudanças na qualidade vocal e na articulação com escopo

<<crepitante>	>	glotalizado
<<sussurrante>	>	mudança de qualidade vocal como indicado

7. Padrão de transcrição de um excerto a partir de uma situação de interação natural, cotidiana

A seguir nós apresentamos um episódio de narrativa conversacional em forma de transcrição básica e de transcrição refinada com notação prosódica intralinear e com ortografia padrão modificada. O exemplo vem de uma interação no Corpus do NUCOI.⁵⁷ Nessa interação, quatro brasileiros compartilham suas experiências após um ano de intercâmbio (Ciência sem Fronteiras, Programa de graduação sanduíche) em Münster, na Alemanha. No trecho a seguir, eles discutem o uso dos pronomes de tratamento formal e informal (*Sie* e *Du*, respectivamente) a partir de suas vivências no período em que estiveram na Alemanha.

O episódio narrativo “Du e Sie” em formato de transcrição básica (ortografia padrão modificada).⁵⁸

01 B1: teve_uma vez que a gente tava no [quiosque]
02 B2: [eHE-]
B1: comprano cerVEja,
03 e_aí o::_o diogo comprou uma garrafinha de: smirnoff_ICE,=
04 =pra::;
05 (-) pra:_pra NANda;=né,
06 (-) aí tem diferença de iDAde;=né,=
07 =eu num manjava disso que cerveja é pra dezesseis Anos,
08 e: bebida (.) destilada_é pra deZOItos anos.
09 B4: AH: não sa[bia] [que era assim.]
10 B1: [e aí] [eu (.) como a gente]
11 B2: [é eu também não sabia]
12 B3: [é MESmo.]
B1: [comprou] a garrafa de smirNOFF,
B2: [NÃO;]
13 B1: o:: (.) o cara o cara do balCÃO_assim;
14 perguntou pro DIOgo assim;=né,=
15 =ah: (.) cê é maior de iDAde;
16 ?ah heHE:- °h

⁵⁷ NUCOI – Núcleo de Estudos de Comunicação (Inter-)Cultural em Interação. Disponível em: <www.letas.ufmg.br/nucleos/nucoi/>, acesso em 10/10/2016.

⁵⁸ O exemplo escolhido difere dos exemplos da versão GAT2 em alemão, bem como do exemplo da versão em inglês, devido ao fato de que, no caso dos dados em português, trata-se da transcrição de uma filmagem e não apenas de áudio. Como consequência, foram integradas algumas informações não verbais.

51 B3: [=foi uma coisa asSIM;]
52 B1: [((ri))]
53 B2: [((ri))]
54 B3: tipo para não me res DESrespeitar;=mas tipo;=
55 =eu não me senti numa (.) posição tipo_eu era comprador e
ela_era sa_sei lá;
56 °h a menina do CAIXA e;;
57 <<rindo> tipo> por que que_ela me trataria> dife!REN!te;=
58 =tá ligado;
59 tipo sei LÁ.
60 pra mim não_ia fazer muito sentido.
61 °hh mas eu lembro que ela ela perguntou tipo_AH-
62 POSSo pergun fazer lhe uma pergunta;
63 [do tipo uma COISA assim;=sabe.]
64 B2: [((ri))]
65 B3: bem <<olhando para câmera> formal.>
66 mas (.) (e) (isso/esse) é engraçado mesmo.

O mesmo episódio narrativo “Du e Sie” em formato de transcrição refinada (ortografia padrão modificada):

01 B1: <<all> teve_uma vez que a gente tava> no [quiosque]
02 B2: [eHE-]
B1: comprando cerveja,
03 e_aí o::_o diogo comprou uma <<indicando tamanho> garrafinha
de: smirnoff_ICE,>=
04 =`pra::;
05 (-) ((click com a língua)) pra:_pra Nanda;=né,
06 (-) aí tem diferença de idade;=né,=
07 =eu num manjava disso que cerveja é pra dezesseis anos,
08 e: bebida (.) destilada_é pra dezoito anos.
09 B4: ↑AH: não sabia] [<dim> que era assim.]
10 B1: [e aí] [eu (.) como a gente]
11 B2: [é eu também não sabia]
12 B3: [<<pp> é MESmo.]]

B1: [comprOu] a garrafa de smir↑NOFF,
 B2: [↓NÃO;]
 13 B1: o:: (.) o cara o cara do balCÃO_assim;
 14 <<h, f> perguntou pro ↑`DIOgo assim;=né,>=
 15 =ah: (.) cê é maior de i`DAde;
 16 ?ah heHE:- °h
 17 <<all> ele perguntou com DU_assIm;=né,>=
 18 =<<all>₁ aí_o diogo olhou assim e falou;=>₁
 19 <<acc>₂ e falou que que>₂ <<h>₃ cê tá a!CHAN!do>₃ velho;
 20 ((risos))
 21 B1: [<<h> sou mais velho que voCÊ;>=]
 22 B2: [((ri))]
 23 B1: [=hahahaHA- °h]
 24 B2: [ah;]
 25 B4: [haHA:-]
 26 B1: aí_o diogo responDEU assim com dU <<dim> também;>=
 27 =<<pp, all> (ele/e) falou> não_eu ^SOU pô.
 28 <<rall> eu tenho vinte e quatro> [Anos;=ahahahaha-]
 29 B2: [NOSSa (.) ↑é tipo de] coisa
 [que se alguém fa]lasse comigo eu> ↑não Ia re`parAr.
 30 B4: [↑é:.]
 31 B2: [se alguém me chamou de <<l> DU_ou SIE,>]
 32 B3: [<<h> também ↑↑NÃO ué.>]
 33 B2: e_aí_eu f eu (.) ↑sei `LÁ;
 34 B1: <<all> o diogo achou> estrANho assim <<all> falou> (.) `pOrra;
 35 cara tá falando [<<cresc> desse jeito coMIgo velho;>=heheHE-]
 36 B2: [que leGAL.=heheheHE-]
 37 B3: [mas AH:-]
 38 B1: [((ri))]
 39 B2: [((ri))]
 40 B3: agOra tu falou <<p> uma coisa que eu me> <<cresc> lembREI;>=
 41 =<<acc> ah eu quando_eu> tava passando no: (.) no CAIXa;
 42 uma menIna (.) °h uma menina achou que_eu era ´↑TURco?
 43 [(-) e ela: e Ela_é ↓acho que_era TURca;=enfIm,]
 44 B1: [((acena com a cabeça e ri))]

45 B3: e ela me pergun↑TOU,
46 (.) ah (-) perguntou se_ela po!DI!a me perguntAr:,
47 ((ri)) se: [eu era TURco.]
48 B2: [((ri))]
49 B3: [<<rindo> entenDEU;>=]
50 B2: [((ri))]
51 B3: [=<<p> foi uma coisa asSIM;>]
52 B1: [((ri))]
53 B2: [((ri))]
54 B3: tipo para não me res DESrespeitar;=mas tipo;=
55 =eu não me <<h> ↑↑sentI numa> (.) posição tIpo_eu era
compradOR <<h> e ela_era> ↓sA_sei lá;
56 °h a menina do ´CAIXa e;;
57 <<rindo> tipO> por que que_ela me trataria> dife!REN!te;=
58 =<<l> tá ligado;>
59 <<p> tipo sei LÁ.>
60 pra mim <<h> não_↑ia fazer> muito senTIdo.
61 °hh <<all> mas eu lembro> que Ela ela perguntou tipo_↑AH-
62 <<len> POSSo pergun fazEr lhe uma pergUnta;>
63 [<<pp> do tIpo uma COIsa assim;=sAbe.>]
64 B2: [((ri))]
65 B3: bem <<olhando para câmara> forMAL.>
66 <<p> mas (.) (e) (isso/esse) é engraÇAdo mesmo.>

Referências

AUER, Peter; COUPER-KUHLEN, Elizabeth; MÜLLER, Frank Ernst. *Language in time*. The rhythm and tempo of spoken interaction. Oxford: Oxford University Press, 1999.

BECKMAN, Mary; AYERS ELAM, Gayle. *Guidelines for ToBI labelling*. Versão 3.0. Columbus: The Ohio State University Research Foundation, 1997. Disponível em: <http://www.cs.columbia.edu/~agus/tobi/labelling_guide_v3.pdf>, acesso em novembro de 2016.

BOLINGER, Dwight. Around the edge of language: Intonation. *Harvard Educational Review*, 34, p. 282-293, 1964. Reimpresso em: BOLINGER, Dwight (org.) *Intonation*. Harmondsworth: Penguin, p. 19-29, 1972.

COUPER-KUHLEN, Elizabeth. *An introduction to English Prosody*. Tübingen: Max Niemeyer; London: Edward Arnold, 1986.

COUPER-KUHLEN, Elizabeth. *English speech rhythm*. Form and function in everyday verbal interaction. Amsterdam: Benjamins, 1993.

COUPER-KUHLEN, Elizabeth. Prosody. In: D'HONDT, Sigurd; ÖSTMAN, Jan-Ola; VERSCHUEREN, Jef. (org.) *The pragmatics of interaction*. Handbook of Pragmatics Highlights, vol. 4. Amsterdam: Benjamins, 2009. p. 174-189.

COUPER-KUHLEN, Elizabeth; SELTING, Margret. *Prosody in conversation*. Cambridge: Cambridge University Press, 1996.

CROFT, William. *Typology and universals*. 2 ed. Cambridge: Cambridge University Press, 2003. p. xix-xxv.

CRUTTENDEN, Alan. *Intonation*. 2 ed. Cambridge: Cambridge University Press, 1997.

DU BOIS, John W.; CUMMING, Susanna; SCHUETZE-COBURN, Stephan; PAOLINO, Danae. Discourse transcription. *Santa Barbara Papers in Linguistics*, Vol. 4, Department of Linguistics, UCSB, 1992.

DU BOIS, John W. (2006): Representing discourse. Transcription in action. Resources for the representation of linguistic interaction. Santa Barbara: Department of Linguistics, University of California, 2006. Disponível em <<http://www.linguistics.ucsb.edu/projects/transcription/>>, acessado em novembro de 2016.

EDWARDS, Jane. Principles and alternative systems in the transcription, coding and mark-up of spoken discourse. In: LEECH, Geoffrey; MYERS, Greg; THOMAS, Jenny. (org.) *Spoken English on computer: Transcription, markup and application*. Harlow: Longman, 1995. p. 19-34.

EDWARDS, Jane A.; LAMPERT, Martin D. *Talking data*. Transcription and coding in discourse research. Hillsdale: Lawrence Erlbaum, 1993.

FERREIRA, Jesuelem S. *O apagamento do /d/ em morfema de gerúndio no dialeto de São José do Rio Preto*. 2010. 142f. Dissertação (Mestrado em Estudos Linguísticos) – Universidade Estadual Paulista Júlio de Mesquita Filho, São José do Rio Preto, 2010.

FORD, Cecilia E.; COUPER-KUHLEN, Elizabeth. Conversation and phonetics: Essential connections. In: COUPER-KUHLEN, Elizabeth; FORD, Cecilia E. (org.) *Sound patterns in interaction*. Cross-linguistic studies from conversation. Amsterdam: Benjamins, 2004. p. 3-25.

FOX, Anthony. *Prosodic features and prosodic structure*. Oxford: Oxford University Press, 2000.

FROTA, Sónia; CRUZ, Marisa; SVARTMAN, Flaviane; COLLISCHONN, Gisela; FONSECA, Aline; SERRA, Carolina; OLIVEIRA, Pedro; VIGÁRIO, Marina. Intonational variation in Portuguese: European and Brazilian varieties In: FROTA, Sónia; PRIETO, Pilar. (org.) *Intonation in Romance*. Oxford: Oxford University Press, 2015a. p. 235–283.

FROTA, Sónia; OLIVEIRA, Pedro; CRUZ, Marisa; VIGÁRIO, Marina. *P-ToBI: tools for the transcription of Portuguese prosody*. Lisboa: Laboratório de Fonética, CLUL/FLUL, 2015b. Disponível em: <<http://labfon.letras.ulisboa.pt/InAPoP/P-ToBI/>>, acessado em novembro de 2016.

GILLES, Peter. *Regionale Prosodie im Deutschen*. Variabilität in der Intonation von Abschluss und Weiterweisung. Berlin; New York: de Gruyter, 2005.

GOODWIN, Charles. Interactive footing. In: HOLT, Elizabeth; CLIFT, Rebecca. (org.) *Reporting talk: Reported speech in interaction*. Cambridge: Cambridge University Press, 2007a. p. 16-46.

GOODWIN, Charles. Participation, stance, and affect in the organization of activities. *Discourse and Society*, 18, p. 53-73, 2007b.

GOODWIN, Charles. Constructing meaning through prosody in aphasia. In: BARTH-WEINGARTEN, Dagmar; REBER, Elisabeth; SELTING, Margret. (org.) *Prosody in interaction*. Amsterdam: Benjamins, 2010. p. 373-394.

GRABE, Esther; NOLAN, Francis; FARRAR, Kimberley. IViE - a Comparative transcription system for intonational variation in English. The 5th International Conference on Spoken Language Processing (ICSLP-1998), 1998, Sydney. *Anais*. p. 1259-1262. Disponível em <http://www.isca-speech.org/archive/icslp_1998/i98_0099.html>, acessado em novembro de 2016.

GRICE, Martine. Intonation. In: Brown, Keith. (org.) *Encyclopedia of Language and Linguistics*. vol. 5. 2 ed. Elsevier: Oxford, 2006. p. 778-788.

GUSSENHOVEN, Carlos. *The phonology of tone and intonation*. Cambridge: Cambridge University Press, 2004.

HEATH, Christian; HINDMARSH, Jon; LUFF, Paul. *Video in qualitative research*. London: Sage, 2010.

JEFFERSON, Gail. Glossary of transcript symbols with an introduction. In: LERNER, Gene. (org.) *Conversation Analysis*. Studies from the first generation. Amsterdam: Benjamins, 2004. p. 13-31.

KERN, Friederike. *Das Zusammenspiel von Prosodie und Syntax am Beispiel von Türkendeutsch*. Habilitationsschrift, Universität Potsdam, 2008.

LADD, D. Robert. *Intonational phonology*. 2 ed. Cambridge: Cambridge University Press, 2008.

LAVER, John. *The phonetic description of voice quality*. Cambridge: Cambridge University Press, 1980.

LEIPZIG GLOSSING RULES. Conventions for interlinear morpheme-by-morpheme glosses. Leipzig, 2008. Disponível em: <<https://www.eva.mpg.de/lingua/pdf/Glossing-Rules.pdf>>, acessado em novembro de 2016.

MONDADA, Lorenza. Multimodal resources for turn-taking: Pointing and the emergence of possible next speakers. *Discourse Studies*, 9, p. 195-226, 2007a. Reimpresso em DIJK, Teun van. (ed.) *Discourse Studies*, 4. London: Sage, 2007. p. 126-157.

MONDADA, Lorenza. Commentary: Transcript variations and the indexicality of transcribing practices. *Discourse Studies*, 9, p. 809-821, 2007b.

MONDADA, Lorenza. Doing video for a sequential and multimodal analysis of social interaction: Videotaping institutional telephone calls. *Forum: Qualitative Sozialforschung; Forum: Qualitative Social Research*, 9, 3, 2008. Disponível em: <www.qualitative-research.net/>, acessado em novembro de 2016.

NESPOR, Marina; VOGEL, Irene. *Prosodic phonology*. 2 ed. Berlin; New York: de Gruyter, 2007.

SCHEGLOFF, Emanuel A. *Sequence organization in interaction*. Cambridge: Cambridge University Press, 2007. Appendix 1, p. 265-269.

SELKIRK, Elisabeth. *Phonology and syntax*. The relation between sound and structure. Cambridge: MIT Press, 1984.

SELTING, Margret. *Prosodie im Gespräch*. Aspekte einer interaktionalen Phonologie der Konversation. Tübingen: Niemeyer, 1995.

SELTING, Margret. Prosody in interaction: State of the art. In: Barth-Weingarten, Dagmar; Reber, Elisabeth; Selting, Margret. (org.) *Prosody in interaction*. Amsterdam: Benjamins, 2010. p. 3-40.

SELTING, Margret; AUER, Peter; BARDEN, Birgit; BERGMANN, Jörg; COUPER-KUHLEN, Elizabeth; GÜNTNER, Susanne; QUASTHOFF, Uta; MEIER, Christoph; SCHLOBINSKI, Peter; UHMANN, Susanne. Gesprächsanalytisches Transkriptionssystem (GAT). *Linguistische Berichte*, 173, p. 91-122, 1998.

SELTING, Margret; AUER, Peter; BARTH-WEINGARTEN, Dagmar; BERGMANN, Jörg; BERGMANN, Pia; BIRKNER, Karin; COUPER-KUHLEN, Elizabeth; DEPPERMANN, Arnulf; GILLES, Peter; GÜNTNER, Susanne; HARTUNG, Martin; KERN, Friederike; MERTZLUFFT, Christine; MEYER, Christian; MOREK, Miriam; OBERZAUCHER, Frank; PETERS, Jörg; QUASTHOFF, Uta; SCHÜTTE, Wilfried; STUKENBROCK, Anja; UHMANN, Susanne: Gesprächsanalytisches Transkriptionssystem 2 (GAT 2). *Gesprächsforschung* 10, p. 353-402, 2009. Disponível em: <www.gespraechsforschung-ozs.de>, acessado em novembro de 2016.

SIMÕES, Adriana M. *Clítico, objeto nulo ou pronome tônico?* Quanto e como a variação/mudança no paradigma do preenchimento pronominal do objeto acusativo de 3ª pessoa no português brasileiro se reflete na aquisição/aprendizagem do espanhol pelos aprendizes brasileiros ao longo das gerações. Dissertação de Mestrado. DLM/FFLCH/USP, São Paulo, 2010.

SORJONEN, Marja-Leena. Simple answers to polar questions: The case of Finnish. In: SELTING, Margret; COUPER-KUHLEN, Elizabeth. (org.) *Studies in Interactional Linguistics*. Amsterdam: Benjamins, 2001. p. 405-431.

STREECK, Jürgen. The body taken for granted: Lingering dualism in research on social interaction. In: GLENN, Phillip; LEBARON, Curtis D.; MANDELBAUM, Jenny S. (org.) *Studies in Language and Social Interaction*. In Honor of Robert Hopper. Mahwah, N.J.: Lawrence Erlbaum, 2003. p. 427-440.

STREECK, Jürgen. *Gesturecraft*. The manufacture of meaning. Amsterdam: Benjamins, 2009.

UHMANN, Susanne. *Fokusphonologie*. Eine Analyse deutscher Intonationskonturen im Rahmen der nicht-linearen Phonologie. Tübingen: Niemeyer, 1991.

VIEIRA, Silvia R. O parâmetro da cliticização fonológica e os pronomes átonos no Português do Brasil e no Português Europeu. *Revista Estudos Linguísticos XXXIV*, p. 1003-1008, 2005.

Data de envio: 22/02/2016
Data de aceite: 26/03/2016
Data da publicação: 23/12/2016