

INTONATION PATTERNS OF TELEPHONE NUMBERS IN BRAZILIAN PORTUGUESE

PADRÕES ENTOACIONAIS DOS NÚMEROS TELEFÔNICOS NO PORTUGUÊS BRASILEIRO

Oyedeji MUSILYU

Faculdade de Letras da Universidade Federal de Alagoas (FALE UFAL)

Miguel OLIVEIRA JR.

Faculdade de Letras da Universidade Federal de Alagoas (FALE UFAL)

RESUMO

O objetivo do presente trabalho é descrever padrões entoacionais de agrupamentos numéricos associados a números telefônicos enunciados em português brasileiro (PB). Para isso, utilizamos uma lista de 30 exemplares de números telefônicos reais, que foram lidos por 85 falantes nativos do PB. Esses enunciados foram segmentados em unidades entoacionais. Os padrões de organização numérica, baseados nessas segmentações, foram então descritos. Em seguida, as variações de frequência fundamental dos enunciados foram analisadas com o auxílio dos scripts MOMEL/INTSINT (Hirst, 2007) e ProsodyPro (Xu, 2012) para o aplicativo Praat (Boersma e Weenink 2013), mediante procedimento semiautomático. Os resultados dessas análises permitiram descrever padrões entoacionais que caracterizam prosodicamente três diferentes tipos de números telefônicos no português brasileiro.

ABSTRACT

The purpose of this study is to identify intonational patterns of numeric groupings associated with spoken telephone numbers in Brazilian Portuguese (BP). To that aim, we use a list of 30 samples of real telephone numbers read aloud by 85 native speakers of BP. The spoken

numbers were segmented into intonation units. The numeric groupings' patterns, based on those segmentations, were then described. After, the variations of fundamental frequency of the spoken numbers were analyzed using MOMEL/INTSINT (Hirst, 2007) and ProsodyPro (Xu, 2012) scripts for Praat (Boersma and Weenink 2013), through a semi-automatic analysis. With the results of those analyses, we were able to describe patterns of intonation that prosodically characterize three different types of telephone numbers in BP.

PALAVRAS-CHAVE

Números telefônicos. Entoação. Português brasileiro.

KEYWORDS

Telephone numbers. Intonation. Brazilian Portuguese.

Introdução

Nos últimos anos, o desenvolvimento da tecnologia da fala tornou o uso de sistemas automatizados de reconhecimento e de síntese de voz bastante frequente, com aplicações múltiplas. Diversos serviços baseados nesses sistemas automatizados fazem uso de agrupamentos numéricos para os mais variados fins, tais como ativação de cartões de crédito, informações bancárias, consultas a listas telefônicas, marcações de reservas, serviços de auxílio a deficientes visuais, entre outros.

Em muitos casos, entretanto, a performance desses sistemas é considerada insatisfatória, ora por não processarem corretamente a fala espontânea (no caso de sistemas de reconhecimento de fala), ora por não apresentarem, em sua produção, características de ritmo e entoação próximas às da fala natural (no caso dos sistemas de síntese de fala). Isso se deve em parte ao fato de serem tais sistemas muitas vezes baseados em dados artificiais e/ou impressionísticos, e não em dados de fala natural e espontânea.

Serviços que fazem uso de números conectados como dados de entrada dependem de um bom sistema de processamento de informações

de unidades sonoras e textuais correspondentes a dígitos alfanuméricos. Um sistema eficiente deverá produzir fala sintética e agrupamentos numéricos correspondentes à enunciação espontânea característica do agrupamento alvo. Sabe-se, todavia, que em agrupamentos numéricos de estrutura fixa (como número telefônico, número de cartão de crédito, número de documento de identidade, etc.), a enunciação de um algarismo vai variar de acordo com a posição que ele ocupa na sequência do agrupamento. Essa variação, em geral, diz respeito aos parâmetros de duração, intensidade e frequência fundamental (Vagner e Sergio, 2005:2). Assim, os parâmetros prosódicos na elocução de um dígito, em uma sequência numérica, vão depender de uma série de fatores, entre os quais a posição que o dígito ocupa na sequência. De acordo com Vagner e Sergio (2005:2), para se obter uma enunciação sintetizada próxima à natural, condizente à pronúncia de um dado dígito em uma determinada sequência numérica, é necessário armazenar no sistema automatizado o maior número possível de informações acerca dos parâmetros prosódicos destes números, no maior número de possibilidades, correspondentes à sua ocorrência em posições sequenciais variadas.

Existe, desse modo, uma premente necessidade de descrever o mais exaustivamente possível as características prosódicas associadas a esses agrupamentos, de forma a contribuir para o aperfeiçoamento desses sistemas computacionais ora em uso. Para isso, é preciso realizar estudos descritivos robustos. Avanços nessa área foram obtidos devido à descrição de números naturais em diversas línguas, tais como o japonês (Amino e Osanai, 2011), o alemão (Baumann e Trouvain, 2001) e o francês (Bartkova e Jouviet, 1999).

Porquanto não exista ainda para o português brasileiro (doravante PB) estudo que descreva, de forma sistemática e abrangente, as características prosódicas da organização de números naturais em estruturas pré-estabelecidas, este estudo apresenta-se como uma contribuição. Como objetivo principal, descreveremos contornos entoacionais típicos de agrupamentos numéricos associados a números

telefônicos enunciados no PB. Para este fim, números telefônicos de diferentes tamanhos serão analisados em duas etapas: primeiramente, serão segmentados em unidades entoacionais, o que permitirá definir as estratégias de agrupamento mais recorrentes. Em seguida, os contornos entoacionais típicos dessas unidades serão analisados e descritos.

1. Estado da arte

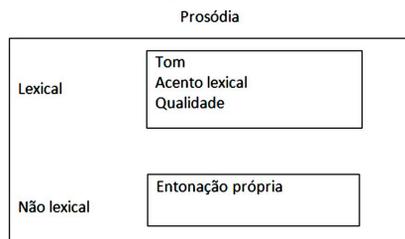
1.1 Prosódia e entoação

O termo “prosódia” tem sido comumente utilizado com equivalente a “entoação” na literatura. Assim, há muita divergência entre os autores, principalmente no que se refere à distinção entre a definição dos dois termos.

A entoação, de acordo com Lidiiane (2007:21 apud Kent e Read, 1992), é parte da prosódia. Ela se refere a um fenômeno mais estreito relacionado geralmente às variações melódicas, enquanto que a prosódia envolve outros parâmetros tais como tempo (pausa e prolongamento) e ritmo.

Hirst e Di Cristo (1998:4), tal como indicado na Figura 1, observam que, o termo prosódia, como o suprasegmental, se relaciona a um sentido amplo diferentemente da entoação que é restringida para o que as vezes é chamado supra lexical, pós-lexical, ou simplesmente características não lexical.

FIGURA 1: Relação entre a prosódia e a entoação



Adaptado de Hirst e Di Cristo (1998:4)

Couper-Kuhlen (1986:71) também apresenta essa distinção afirmando que a entoação no sentido restrito da melodia do discurso é visto neste modelo como uma manifestação da interação de várias características prosódicas, em primeiro lugar o tom (direção do *pitch*) e *pitch-range* (altura e largura do *pitch*). Outras características prosódicas, como sonoridade, ritmo, pausa, etc, podem, no entanto, também entrar nesse caso, estamos falando de entoação em seu sentido mais amplo, ou prosódia.

O termo entoação, neste estudo, se refere ao uso restrito da palavra, como sendo a parte da prosódia que diz respeito à variação da frequência fundamental na fala (variação melódica) (Kent e Read, 1992).

1.2 Síntese de frequência fundamental (F0)

Nos últimos anos, vários estudos têm se voltado para contribuir com o desenvolvimento de *software* amplamente acessível para a manipulação e análise de texto e fala, como, por exemplo, aquele desenvolvido pelo projeto europeu *MULTEXT* (Véronis et al., 1994). Um dos objetivos desses estudos é a **integração de algumas ferramentas e métodos de análise da fala natural e da linguagem através da prosódia.**

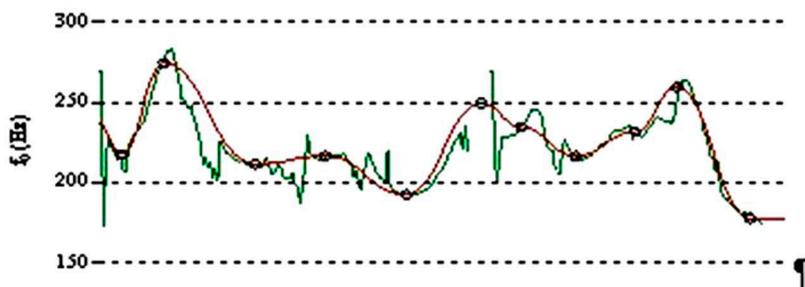
Nessa veia, alguns laboratórios desenvolveram e aplicaram a extração automática da informação entonativa da fala pelo processo de estilização (Campioni et al., 2000). A tarefa é desempenhada por ferramentas tal como o *MOMEL*, que permite a estilização automática de F0 como pontos-alvo (Hirst e Espesser, 1993).

O algoritmo *MOMEL* permite a representação de F0 através duma sequência de pontos-alvo constituídos por pares de valores <F0, tempo>. Os pontos-alvo correspondem às variações locais pertinentes da curva melódica e permitem, a partir de uma função quadrática de tipo *spline*¹, recuperar o perfil suprasegmental, caracterizando de modo geral a entoação. No trabalho do Hirst (2005), através dessa função, observa-

¹ Um *spline* é uma curva definida matematicamente por dois ou mais pontos de controle.

se que a estilização produz uma curva bem próxima à curva original de frequência fundamental, sem perda de informações significativas, tal como exemplificado na Figura 2:

FIGURA 2: A janela do praat com a curva entoacional realizada pelo MOMEL, em vermelho, do número telefônico 2226 3196 tal como falado pela participante Mulher_03.



1.3 Sistema de anotação prosódica da fala: INTSINT (International Transcription System for Intonation)

Vários modelos de anotação prosódica foram desenvolvidos e aplicados no estudo da entoação. O modelo aqui descrito é o *INTSINT* (Hirst e Di Cristo, 1998). Trata-se de um modelo de base fonético-fonológica que, ao contrário de muitos outros, desenvolvidos para uma língua específica, objetiva, aos modos do Alfabeto Fonético Internacional (IPA), ser um sistema internacional de transcrição para entoação, como sua própria sigla sugere (*IN*ternational *T*ranscription *S*ystem for *IN*Tonation). A transcrição a partir deste modelo é obtida, de forma semi-automática, através do *script* MOMEL/INTSINT, quando utilizado em conjunto com o *software Praat* (Boersma e Weenink, 2013).

Para a descrição da entoação, é usada uma série limitada de símbolos tonais associados aos pontos-alvo estimados pela técnica MOMEL (*Modelisation de Melodie*) (Louw e Barnard, 2004), descrita acima. Tais

símbolos ortográficos abstratos definidos para representar esses pontos-alvo estão listados no Quadro 1, seguidos da sua abreviatura e da proposta de tradução para o português:

QUADRO 1: Configurações de acentos tonais de acordo com o INTSINT.

Top	T	Topo
Mid	M	Médio
Bottom	B	Base
higher	H	Mais alto
Same	S	Igual
Lower	L	Mais baixo
Upstepped	U	Subida suave
Downstepped	D	Descida suave

1.4 Propriedades dos números telefônicos no Brasil

Os números telefônicos são geralmente agrupados, graficamente, em blocos de um, dois, três ou quatro dígitos. O enunciado do número 32214754 pode, por exemplo, ser representado como a seguir:

3221-4754 agrupamento de forma (NNNN-NNNN)² ou
 3221-47-54 agrupamento de forma (NNNN-NN-NN) ou
 3-221-47-54 agrupamento de forma (N-NNN-NN-NN)³

No Brasil, os números de telefone são regulados pela ANATEL (Agência Nacional de Telecomunicações), no Plano Nacional de Numeração. No plano atual, o número de assinante deverá ter o comprimento de oito dígitos, obedecendo ao formato [NNNN + NNNN]⁴. Os números especiais de ligação gratuita, geralmente,

² A letra “N” aqui representa um dígito.

³ Os hifens são opcionais. Podem ser substituídos por espaços em branco.

⁴ Segundo uma recomendação nova, que só se aplica até o momento a telefones móveis, em algumas regiões do Brasil os números de telefones deverão ter o comprimento de 9 (nove) dígitos, obedecendo ao formato [9NNNN + NNNN].

apresentam um comprimento de três ou onze dígitos, obedecendo ao formato [NNN] e [NNNN + NNNNNNNN] respectivamente. São números telefônicos de entidades prestadoras de serviços de socorro e de informações de interesse comunitário, por exemplo, o 190 (polícia militar) e o 0800 5700100 (correio).

É importante notar que existem no Brasil, além dos números telefônicos propriamente ditos, prefixos numéricos opcionais que são utilizados em ligações especiais, como é o caso do código DDD (discagem direta a distância), associado a cada estado brasileiro, e o prefixo numérico associado a diferentes operadoras telefônicas. O “DDD” é uma numeração de códigos de dois dígitos criada para fazer ligações interestaduais. O código da operadora é um conjunto de dois dígitos únicos para cada operadora. No presente estudo, esses prefixos não foram considerados.

1.5 Estrutura prosódica dos números telefônicos

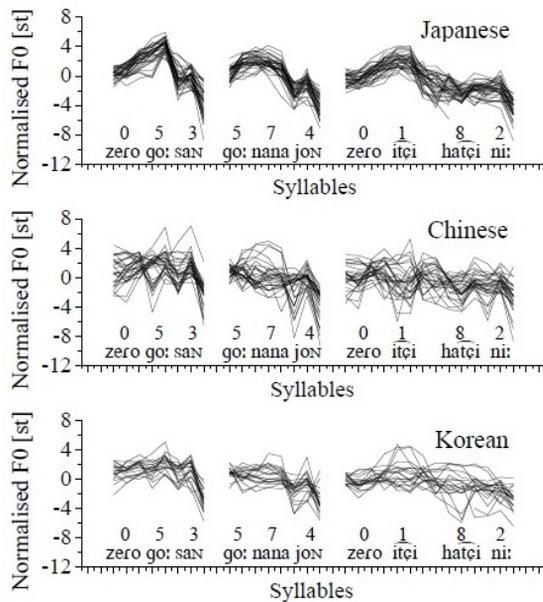
Foi apenas nos últimos anos, com o advento da tecnologia de fala, que a prosódia de expressões recorrentes pré-definidas (tais como horários, informações de conta bancária, de cartões de crédito, de listas telefônicas etc...) começou a receber uma atenção adequada. Isso se deve principalmente à necessidade de um aperfeiçoamento dos sistemas computacionais, que precisam processar corretamente a fala, produzindo unidades sonoras com uma prosódia **próxima à da fala natural** (Hakulinen et al., 1999).

De acordo com a literatura, os números de telefone apresentam uma prosódia particular, que pode ser descrita em termos de melodia e ritmo na fala (Baumann e Trouvain, 2001). A seguir, resultados de trabalhos feitos para as línguas japonesa, francesa e alemã serão apresentados sucintamente, com o objetivo de ilustrar como a prosódia de números telefônicos tem sido descrita em diferentes línguas.

Amino e Osanai (2011) realizaram um estudo sobre a estrutura

prosódica na enunciação dos números de telefone japoneses por falantes nativos e não nativos do japonês. Os autores observaram que no japonês os números de telefone têm uma estrutura prosódica bipartida: na sua enunciação, o acento tonal ocorre a cada dois dígitos. Eles analisaram e compararam os contornos entoacionais na elocução dos números de telefone por japoneses nativos e aprendizes do japonês chineses e coreanos. Os resultados revelaram que só os falantes nativos reproduziram a estrutura entoacional de modelo bipartido e os falantes não nativos apresentaram uma estrutura prosódica diferente, a depender das suas línguas nativas. Na Figura 3, os resultados são exemplificados com as variações da F0 entre os falantes nativos e não nativos do japonês:

FIGURA 3: O contorno da F0 (em semitons) do número 053 574 0182 enunciado por falantes nativos do japonês (superior), chinês (meio) e coreano (inferior).

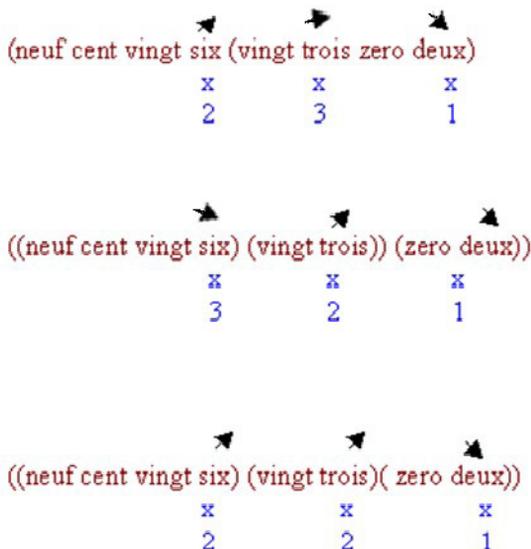


Fonte: Amino e Osanai (2011:4)

Martin (1997) observou que na enunciação dos números de telefone em francês, as variações do contorno melódico não têm obrigatoriamente relação semântica, mas são relacionadas às escolhas de segmentações, que podem ser individuais ou culturais. Assim, na França, o número de telefone “9262302”, por exemplo, pode ser segmentado em pelo menos três maneiras diferentes: (926)(23-02), (926-23)(02) e (926)(23(02)).

Cada segmentação incorrerá em contornos entoacionais específicos. A Figura 4 exemplifica as variações ascendentes (setas para cima) e descendentes (setas para baixo) do contorno entoacional dessas segmentações:

FIGURA 4: O contorno entoacional de diferentes tipos de enunciação do número telefônico 9262302 na França.



Fonte: Martin (1997)

Martin (1997) observou também que o contorno entoacional na enunciação dos números de telefone varia em relação à modalidade enunciativa. Assim, em uma enunciação interrogativa, a última

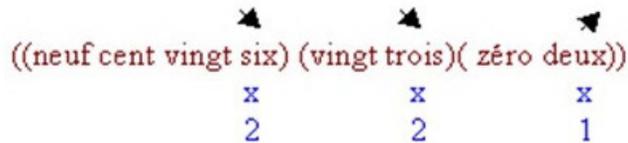
segmentação termina com um contorno ascendente e os segmentos anteriores terminam com um contorno descendente. Numa enunciação afirmativa, a última segmentação termina com um contorno descendente e os segmentos anteriores terminam com um contorno ascendente, tal como exemplificado respectivamente nas Figuras 5 e 6:

FIGURA 5: O contorno entoacional interrogativo de um tipo de enunciação do número telefônico 9262302 na França.



Fonte: Martin (1997)

FIGURA 6: O contorno entoacional afirmativo de um tipo de enunciação do número telefônico 9262302 na França.



Fonte: Martin (1997)

Baumann e Trouvain (2001) realizaram um estudo sobre a estrutura prosódica dos números de telefone no alemão. Os autores fizeram uma análise de produção para investigar a estratégia de agrupamento e de enunciação dos números de telefone e em seguida, realizaram uma análise

experimental de percepção para testar em serviços de informações telefônicas a estratégia de produção mais comumente observada.

Os autores afirmam que não há diferença entre os dígitos em termos de seus valores semânticos, mas sim há diferença no que diz respeito a seus valores prosódicos. Nas suas análises dos aspectos prosódicos, Baumann e Trouvain (2001) adotaram os pressupostos da Fonologia Auto-segmental Métrica (Ladd, 1996) baseados numa relação expressa como *weak-strong pattern*, ou seja, um padrão de alterações de sílabas fracas e fortes. Nesta relação, a estrutura tonal é sobreposta sobre a estrutura métrica, o que significa que o elemento forte (S, do inglês *strong*) vai receber um *pitch accent* ou acento tonal, enquanto o elemento fraco (W, do inglês *weak*) não vai.

O Quadro 2 e a Figura 7 exemplificam respectivamente o modelo de produção e de estrutura prosódica dos números de telefone no alemão, tal como descrito por Baumann e Trouvain (2001). As estruturas prosódicas padrões observadas nas análises de produção são ilustradas na Figura 7 e são descritas pelo modelo de notação entoacional *GtoBF*⁵ baseado na Fonologia Auto-segmental Métrica:

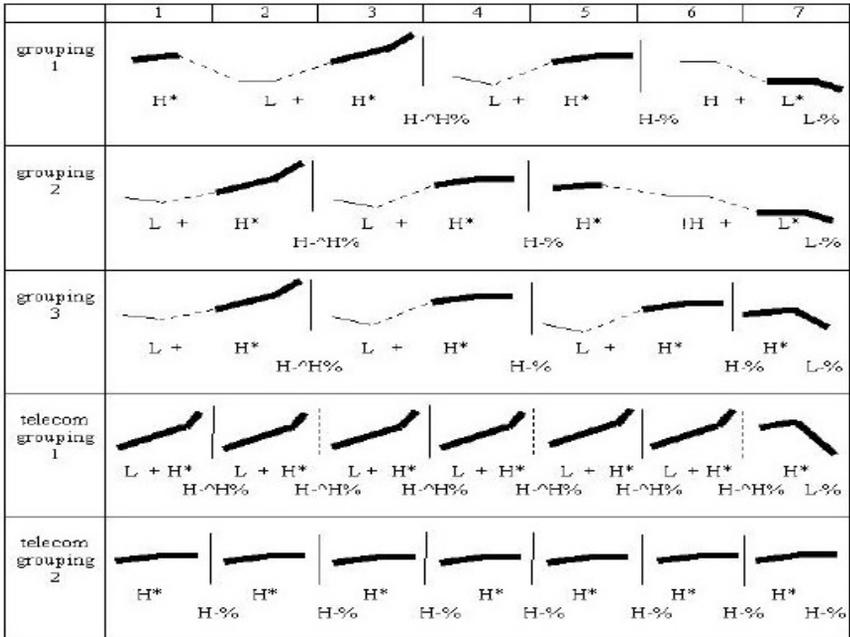
QUADRO 2: As estratégias com agrupamentos e estruturas métricas

strategy	grouping	metrical structure
NAT 1	3-2-2	SWS - WS - WS
NAT 2	2-2-3	WS - WS - SWS
NAT 3	2-2-2-1	WS - WS - WS - S
TEL 1	1-2-2-2	S - S(-)S - S(-)S - S(-)S
TEL 2	1-1-1-1-1-1-1	S - S - S - S - S - S - S

Fonte: Baumann e Trouvain (2001:2)

⁵ Sistema fonológico de transcrição da entonação do alemão.

FIGURA 7: contornos esquemáticos. As linhas representam: acentuada (forte), átona (fina), transição (pontilhada), IP-fronteiras (vertical escura), ip-fronteiras (vertical pontilhada)



Fonte: Baumann e Trouvain (2001:2)

2. Metodologia

2.1 A caracterização do corpus

O *corpus* deste estudo, como apresentado abaixo, no Quadro 3, consiste de um total de 30 números telefônicos. Todos os números são reais e foram extraídos da lista telefônica LISTEL das cidades de Recife, Olinda e Jaboatão dos Guararapes, no Estado de Pernambuco:

QUADRO 3: O corpus do estudo.

Números telefônicos			
3 dígitos	8 dígitos		11 dígitos
120	32224034	2226 31 96	08002812112
104	32514251	3221 47 54	08007010114
147	33274686	3271 00 84	08007011566
190	34238577	3428 09 24	08007070044
193	34412276	3465 30 46	08007704418
	3228 6924	8803 91 48	
	3251 7343	9605 36 81	
	3424 2767	9619 94 53	
	3452 1425	9909 62 94	
	3465 2746	9948 09 93	

Os números foram escolhidos aleatoriamente, de forma a abranger (a) os números convencionais com oito dígitos, (b) os números de telefone celular com oito dígitos, (c) os números de serviços especiais com três dígitos e (d) os números de ligação gratuita com onze dígitos.

Procurou-se, na medida do possível, utilizar combinações de números variadas. No intuito de testar uma possível relação entre a apresentação gráfica dos números e a maneira como são enunciados, os números convencionais de oito dígitos foram apresentados de três maneiras diferentes, (i) divididos em dois grupos de quatro dígitos (NNNN NNNN), (ii) divididos em um grupo de quatro e dois grupos de dois dígitos (NNNN NN NN) e (iii) sem qualquer tipo de divisão (NNNNNNNN). Os catálogos telefônicos consultados trazem números com esses três tipos de disposição, embora a do tipo (i) seja a mais frequente.

2.2 Descrição dos participantes

Os participantes que produziram os dados utilizados neste estudo eram, à época da coleta, estudantes (uma metade do segundo grau e outra do terceiro grau) falantes nativos do PB, todos do estado de Pernambuco. Um total de oitenta e cinco pessoas (quarenta e oito mulheres e trinta e sete homens) participou da coleta de dados de forma voluntária.

2.3 Coleta de dados

Na coleta de dados, os números de telefone pré-selecionados foram apresentados um após outro, aleatoriamente, em *slideshow*, com intervalos regulares de sete segundos. O processo dividiu-se em duas etapas para cada participante.

Na primeira etapa, o participante leu uma breve instrução para deixá-los cientes de que eles não estavam sendo testados no processo da coleta de dados, possibilitando assim que sua fala fosse o mais espontânea possível. Depois da leitura da instrução, um rápido ensaio, contendo um total de seis números telefônicos representativos (9619 94 53; 32514251; 104; 0800701 1566; 3228 6924; 8803 91 48) foi realizado.

Na segunda etapa, logo depois do ensaio, o *slideshow* composto dos 30 números foi apresentado aos participantes e a gravação foi imediatamente feita. A gravação foi realizada em um equipamento de *minidisc Sony*, modelo MZ-R700, com um microfone digital *Sony*, modelo ECM-MS907, localizado a cinco centímetros da boca dos participantes.

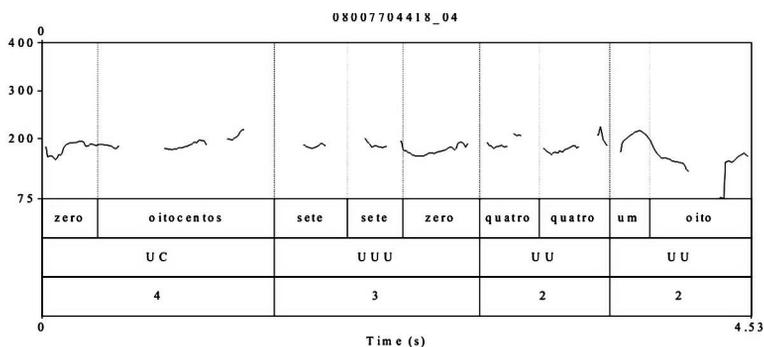
No processo da leitura, sete e doze participantes omitiram respectivamente um e dois números telefônicos. Portanto, para esses participantes, utilizamos apenas dados de 29 e 28 enunciados de números telefônicos respectivamente, em vez de 30 enunciados.

Também, para as análises desse estudo, não foram utilizados todos os dados produzidos pelos participantes. Quarenta e cinco (45) participantes erraram na leitura de um, dois, três, quatro ou sete números telefônicos. Consequentemente, um total de 72 enunciados de números telefônicos foi descartado da análise. No entanto, o número de dados omitidos (31 enunciados) e descartados (72 enunciados) é pouco significativo. Ao todo, foi analisado um total de 2.447 enunciados gravados.

2.4 Análise segmental dos dados

Cada um dos números telefônicos produzidos pelos participantes foi individualmente transcrito e segmentado em palavras e em unidades entoacionais⁶. As unidades entoacionais foram categorizadas no que diz respeito à quantidade de números nelas encapsulados (sendo o mínimo de 1 número e o máximo de 4 números) e ao(s) tipo(s) de unidade matemática, no sistema decimal de numeração, nelas presentes, representados pela letra U (unidade), D (dezena), C (centena) ou M (milhar). A Figura 8 exemplifica a segmentação e transcrição de um número telefônico tal como enunciado por um falante:

FIGURA 8: Janela do *Praat* contendo a curva da F0. No *textgrid*, a transcrição e a segmentação do número 08007704418 tal como enunciado pela participante Mulher_4.



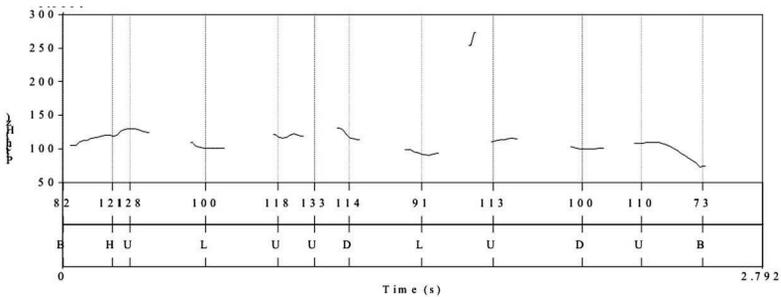
2.5 Anotação entoacional dos dados

Cada número telefônico foi processado semi-automaticamente por meio do *script* *MOMEL/INTSINT for Praat* (Hirst, 2007). Primeiro, através do *MOMEL*, o *script* faz a modelização da curva de F0 e depois a

⁶ Segmento funcional básico do discurso, definido principalmente em termos de contornos entoacionais coesivos, delimitados por elementos prosódicos diversos, isolados ou em conjunto, tais como pausas, mudanças na F0, na duração, na intensidade e na qualidade de voz (CHAFE, 1994).

anotação baseada nessa modelização. A Figura 9 exemplifica a descrição do contorno entoacional de um número tal como enunciado por um falante:

FIGURA 9: Janela do *Praat* contendo a curva da F0. No *textgrid*, os valores da F0 estilizados pelo *MOMEL* e a codificação do *INTSINT* do número telefônico 08007704418 tal como enunciado pelo participante Homem_69.



Como é possível observar na Figura 9, acima dos símbolos de anotação gerados pelo *INTSINT*, estão os valores de variação da F0 em *Hertz* (Hz). Esses valores são pontos-alvo de F0 associados às anotações. Assim, por exemplo, o *script* *MOMEL/INTSINT* considerou o ponto-alvo a 121 Hz como um tom alto (H) e, mais adiante, o ponto-alvo correspondente a 91 Hz como um tom baixo (L).

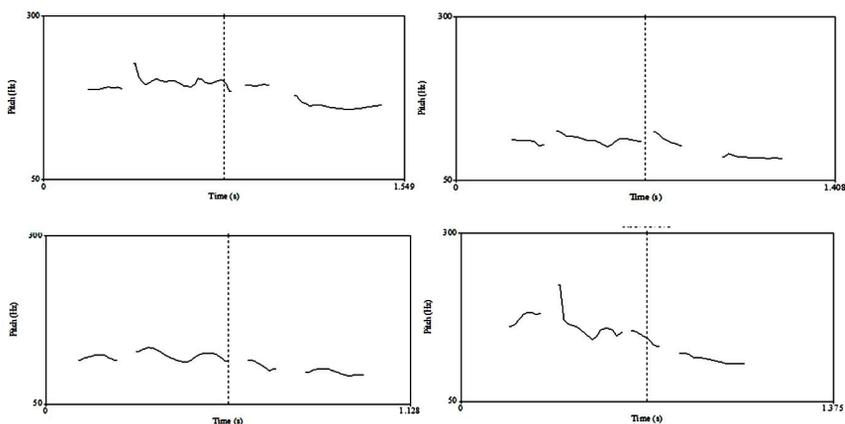
2.6 Representação de contornos entoacionais: *ProsodyPro*

O *ProsodyPro* foi desenvolvido por Xu (1999). É um *script* utilizado em conjunto com o *software Praat* que facilita a análise prosódica em larga escala, principalmente, para dados experimentais. O ponto fundamental do funcionamento desse programa é o alinhamento temporal para facilitar a comparação direta de contornos contínuos de F0, enquanto gera medições múltiplas (adequadas para análise estatística) de dados não alinhados temporalmente. A partir desse método de alinhamento

temporal, o *script ProsodyPro* possibilita um exame minucioso de contornos de F0 contínuos em vários dados.

Os exemplos a seguir mostram como contornos entoacionais podem ser facilmente identificados por alinhamento temporal realizado através do *script ProsodyPro*. Eles foram extraídos das análises do presente estudo e os dados são referentes ao enunciado do número telefônico “193” por quatro participantes. A Figura 10 traz os contornos entoacionais do enunciado do número telefônico “193” pelos participantes Mulher_04, Homem_63, Homem_57 e Homem_78 respectivamente:

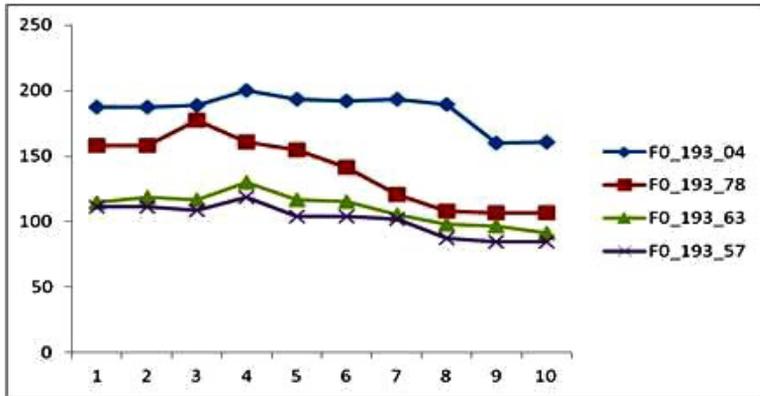
FIGURA 10: Contornos entoacionais do enunciado do número telefônico “193” pelos participantes Mulher_04, Homem_57, Homem_63 e Homem_78.



Na Figura 10, podemos observar, visualmente, algumas semelhanças nos contornos entoacionais dos participantes ao enunciar o número telefônico “193”. Para capturar melhor essas semelhanças, realizamos uma análise com o *script ProsodyPro*. Para cada enunciado, o *script* coleta valores de F0 em intervalos de tempo equidistantes. A quantidade de vezes que o *script* coleta essa informação pode ser predefinida pelo pesquisador (dez vezes neste estudo). A partir dos dados de F0, estilizados com base

nas informações de dez pontos-alvo de F0 coletados pelo *ProsodyPro*, elaboramos um gráfico com os contornos das unidades entoacionais em separado, como ilustrado no Gráfico 1:

GRÁFICO 1: Contornos entoacionais em alinhamento temporal de enunciados do número telefônico “193” por múltiplos participantes.



No Gráfico 1, é possível observar que o uso de pontos-alvo de F0 distribuídos equidistantemente possibilita uma melhor comparação, mediante uma visualização mais adequada, dos contornos entoacionais de um mesmo enunciado produzido por diferentes falantes, com durações diferentes. A partir dessa comparação, pode-se fazer considerações mais adequadas acerca da configuração de um contorno entoacional relacionado a uma mesma unidade linguística produzida por diferentes falantes.

3. Resultados

3.1 A estratégia de agrupamento dos números telefônicos.

Como se pode observar na Tabela 1, a preferência do agrupamento dos números de oito dígitos em sua maioria é no tipo 2-2-2-2, isto é, em

quatro agrupamentos binários (85% dos casos). Em segundo lugar na preferência dos falantes, em apenas 11% dos casos, é o agrupamento do tipo “1-3-2-2” (N NNN NN NN).

TABELA 1: Resultados da estratégia de agrupamento dos números de oito dígitos.

Números de dígitos Agrupamentos	Oito dígitos		
	2-2-2-2	1-3-2-2	Outros (3-2-3; 1-1-2-2-2; ...)
Total	1435	188	70
%	85%	11%	4%

Na Tabela 2, é possível observar que a disposição gráfica dos números de oito dígitos não parece exercer influência na maneira como eles são agrupados pelos participantes. As ocorrências do agrupamento “2-2-2-2” (NN NN NN NN) foram de 85% para os casos de disposição gráfica NNNN NNNN, 80% para NNNNNNNN e 81% dos casos para a disposição gráfica NNNN NN NN:

TABELA 2: Resultados da estratégia do agrupamento com os números de disposição gráfica NNNN NNNN, NNNNNNNN e NNNN NN NN.

Números de dígitos	Oito dígitos		
Disposição gráfica	NNNNNNNN	NNNN NNNN	NNNN NN NN
Total	360	326	749
% de agrupamento 2-2-2-2 (NN NN NN NN)	80%	85%	81%

Como se pode observar na Tabela 3, a preferência do agrupamento dos números de três dígitos em sua totalidade é no tipo “3” (NNN), isto é, um agrupamento ternário.

TABELA 3: Resultados da estratégia de agrupamento dos números de três dígitos.

Números de dígitos	Três dígitos
Agrupamentos	3
Total	424
%	100%

Por fim, a Tabela 4 apresenta os números de onze dígitos, indicando que o agrupamento é majoritariamente feito em “4-3-2-2” (NNNN NNN NN NN), isto é, um conjunto de um agrupamento quaternário, seguido de um agrupamento ternário e dois agrupamentos binários (98% dos casos):

TABELA 4: Resultados da estratégia de agrupamento dos números de onze dígitos.

Números de dígitos	Onze dígitos	
Agrupamentos	4-3-2-2	Outros (2-2-3-2-2; 4-2-1-2-2; ...)
Total	319	5
%	98%	2%

3.2 A estratégia de organização dos números telefônicos em unidades decimais

A seguir, serão descritos os padrões de organização dos números telefônicos em unidades decimais, considerando-se apenas o padrão de agrupamento mais frequente em cada tipo de número telefônico. A Tabela 5 abaixo traz os padrões que foram encontrados para o agrupamento mais frequente em números de oito dígitos:

TABELA 5: Resultados da distribuição sonora dos números telefônicos de oito dígitos de agrupamentos binários “2-2-2-2” (NN NN NN NN).

Números de dígitos	Oito dígitos				
	UU-UU-UU-UU	D-D-D-D	UU-UU-D-UU	UU-UU-D-D	Outros
Total	689	135	135	130	346
%	48%	9%	9%	9%	24%

Como se pode observar na Tabela 5, a preferência de organização dos números telefônicos de oito dígitos de agrupamento “2-2-2-2” (NN NN NN NN), em sua maioria, é pelo padrão “UU-UU-UU-UU” que é uma organização numérica em unidades (48% dos casos). Outros (24% dos casos) são tipos de distribuições sonoras correspondentes a no máximo 4%.

A preferência de organização dos números telefônicos de três dígitos de agrupamento “3” (NNN), como se pode observar na Tabela 6 abaixo, é, por sua vez, pelo padrão “C”, que é uma organização numérica em centena (66% dos casos):

TABELA 6: Resultados da distribuição sonora dos números telefônicos de três dígitos de agrupamento ternário “3” (NNN)

Números de dígitos	Três dígitos	
	UUU	C
Total	144	280
%	34%	66%

Por fim, a Tabela 7 aponta a preferência de organização dos números telefônicos de onze dígitos de agrupamento “4-3-2-2” (NNNN NNN NN NN):

TABELA 7: Resultados da distribuição sonora dos números telefônicos de onze dígitos de conjunto de agrupamentos “4-3-2-2” (NNNN NNN NN NN).

Números de dígitos	Onze dígitos				
Distribuição sonora	UC-UUU-UU-UU	UC-C-D-D	UC-UUU-D-D	UC-C-D-UU	Outros
Total	62	55	52	43	112
%	19%	17%	16%	13%	35%

Neste caso, a preferência recai sobre o padrão “UC-UUU-UU-UU”, que é um agrupamento quaternário de distribuição em unidade e centena, seguido de um agrupamento ternário de distribuição sonora em unidade e de dois agrupamentos binários de distribuição sonora em unidade (19% dos casos). É importante observar aqui que a preferência neste caso é apenas marginal. Este padrão de realização pouco consistente no caso de números telefônicos de onze dígitos pode estar relacionado ao fato de serem números muito pouco usuais, se comparados aos números de oito dígitos e aos de três dígitos.

3.3 Padrões entoacionais dos números telefônicos.

Na enunciação dos números telefônicos de três dígitos de distribuição C, a anotação entoacional mais recorrente gerada pelo script *INTSINT/MOMEL* foi aquela representada pela combinação de acentos tonais M U D, com 39% de ocorrência. Outros (36% dos casos) são tipos de anotações entoacionais correspondentes a no máximo 1%, tal como ilustrado na Tabela 8:

TABELA 8: As anotações recorrentes geradas pelo *script INTSINT/MOMEL* na enunciação dos números de três dígitos de distribuição C, tal como enunciada pelos participantes.

Anotação	MUD	DD	MDUD	MDB	TDDU	TDU	TDUD	Outros
%	39	6	4	2	4	6	4	36

Na enunciação dos números telefônicos de oito dígitos de distribuição UU-UU-UU-UU, a anotação entoacional mais recorrente gerada para a primeira unidade entoacional UU(1)⁷ foi a combinação de acentos M U D, com 39% de ocorrência tal como ilustrado na Tabela 9:

TABELA 9: As anotações recorrentes geradas pelo *script INTSINT/MOMEL* na enunciação de UU(1) dos números de oito dígitos, tal como enunciada pelos participantes.

Anotação	MUD	TD	BDD	BUD	Outros
%	39	9	2	14	14

Para a segunda unidade entoacional UU(2) foi a combinação D U, com 33% de ocorrência tal como ilustrado na Tabela 10:

TABELA 10: As anotações recorrentes geradas pelo *script INTSINT/MOMEL* na enunciação de UU(2) dos números de oito dígitos, tal como enunciada pelos participantes.

Anotação	DU	UDU	DB	UD	D	Outros
%	33	4	2	10	15	36

Para a terceira unidade entoacional UU(3) foi a combinação U D, com 35% de ocorrência tal como ilustrado na Tabela 11:

⁷ Aqui os números entre parênteses serão utilizados para discriminar as unidades entonacionais que possuem características distribucionais e organizacionais idênticas.

TABELA 11: As anotações recorrentes geradas pelo *script INTSINT/MOMEL* na enunciação de UU(3) dos números de oito dígitos, tal como enunciada pelos participantes.

Anotação	UD	DUD	DD	DU	D	Outros
%	35	5	3	5	15	37

Por fim, para a quarta unidade entoacional UU(4) foi a combinação U B, com 36% de ocorrência tal como ilustrado na Tabela 12:

TABELA 12: As anotações recorrentes geradas pelo *script INTSINT/MOMEL* na enunciação de UU(3) dos números de oito dígitos, tal como enunciada pelos participantes:

Anotação	D	UB	DD	DB	B	Outros
%	6	36	2	3	2	50

Na enunciação dos números telefônicos de onze dígitos de distribuição UC-UUU-UU-UU, a anotação entoacional mais recorrente gerada para a primeira unidade entoacional UC foi a combinação de acentos tonais M U D U, com 48% de ocorrência tal como ilustrado na Tabela 13:

TABELA 13: As anotações recorrentes geradas pelo *script INTSINT/MOMEL* na enunciação de UC dos números de onze dígitos, tal como enunciada pelos participantes.

Anotação	MUDU	MUU	MUD	Outros
%	48	5	6	24

Para a segunda unidade entoacional UUU foi a combinação M D, com 37% de ocorrência tal como ilustrado na Tabela 14:

TABELA 14: As anotações recorrentes geradas pelo *script INTSINT/MOMEL* na enunciação de UUU dos números de onze dígitos, tal como enunciada pelos participantes.

Anotação	TDD	MT	DD	MD	Outros
%	3	5	6	37	48

Para a terceira unidade entoacional UU(1) foi a combinação U D, com 50% de ocorrência tal como ilustrado na Tabela 15:

TABELA 15: As anotações recorrentes geradas pelo *script INTSINT/MOMEL* na enunciação de UU(1) dos números de onze dígitos, tal como enunciada pelos participantes.

Anotação	DU	D	U	UD	Outros
%	5	8	3	50	34

Para a quarta unidade entoacional UU(2) foi a combinação U B, com 44% de ocorrência tal como ilustrado na Tabela 16:

TABELA 16: As anotações recorrentes geradas pelo *script INTSINT/MOMEL* na enunciação de UU(2) dos números de onze dígitos, tal como enunciada pelos participantes.

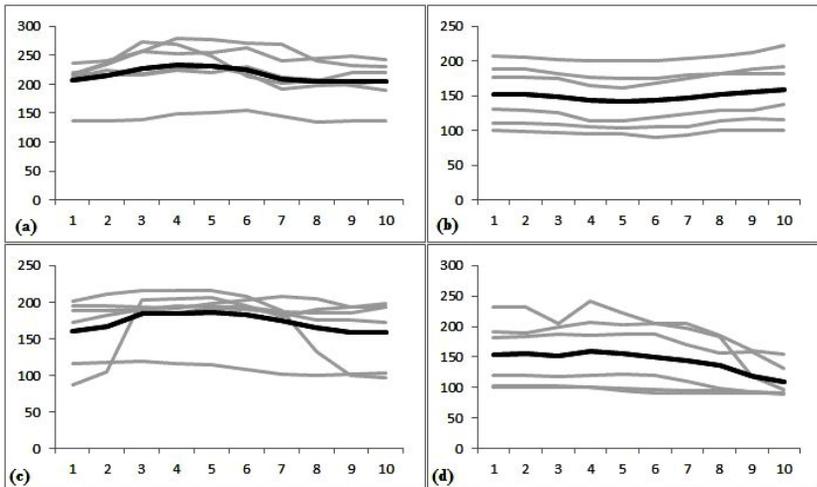
Anotação	MD	DU	DD	UB	Outros
%	3	3	8	44	42

3.4 Os contornos entoacionais dos números telefônicos

Analizamos a variação de valores de F0 selecionados pelo *ProsodyPro* em cada unidade entoacional de cada um dos números telefônicos investigados que seguiram os padrões de agrupamento e organização identificados na primeira etapa deste estudo.

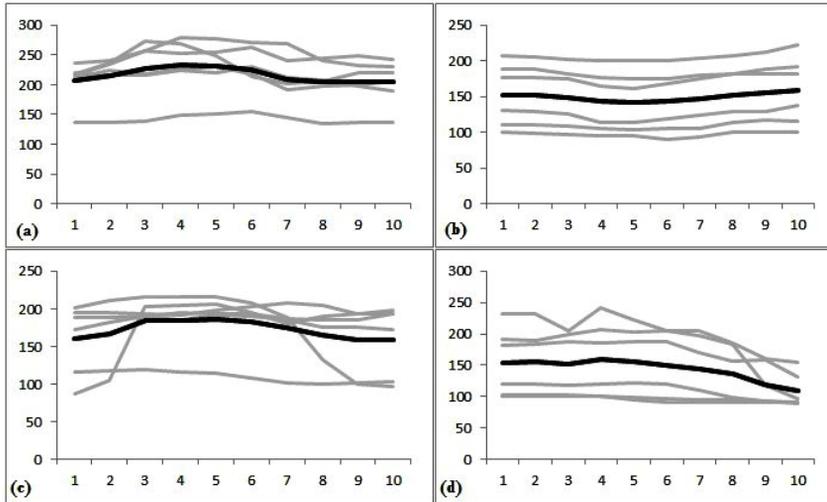
Acerca dos números telefônicos de oito dígitos, o Gráfico 2 contém os contornos individuais e modelizados das unidades entoacionais UU(1), UU(2), UU(3) e UU(4), respectivamente ilustrados nos gráficos (a), (b), (c) e (d), tal como falado por diferente participantes:

GRÁFICO 2: Os contornos entoacionais em unidades entoacionais UU(1), UU(2), UU(3) e UU(4) e seus respectivos padrões modelizados.



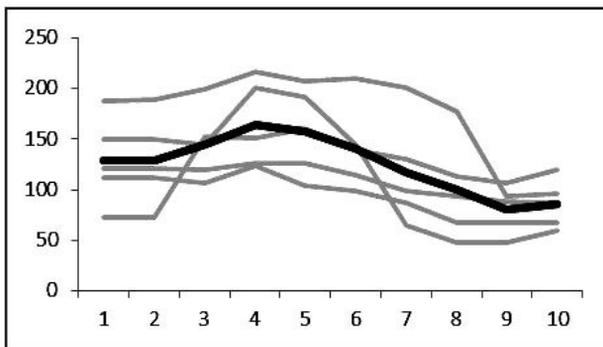
O Gráfico 3 traz informações acerca dos contornos das unidades entoacionais UC, UUU, UU(1) e UU(2), respectivamente observados nos gráficos (a), (b), (c) e (d), bem como dos padrões modalizados de cada uma delas para números de onze dígitos:

GRÁFICO 3: Os contornos entoacionais em unidades entoacionais UC, UUU, UU(1) e UU(2) e seus respectivos padrões modelizados.



Por fim, o Gráfico 4 ilustra os contornos entoacionais individuais da unidade entoacional C, típica de números telefônicos de três dígitos, bem como o seu padrão modelizado:

GRÁFICO 4: Os contornos entoacionais em unidades entoacionais C e seu respectivo padrão modelizado.



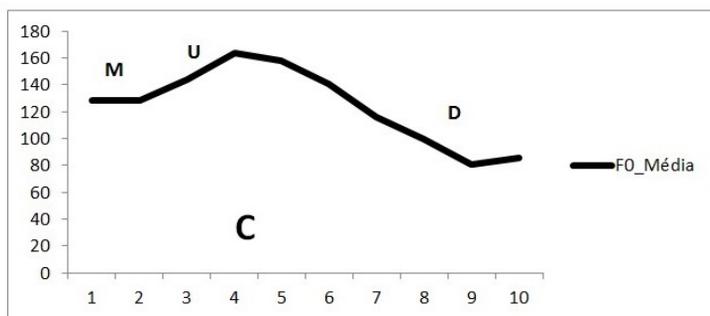
4. Discussão

O objetivo principal da análise aqui proposta foi descrever os padrões entoacionais relacionados à enunciação de números telefônicos do português falado no Brasil.

Consideramos primeiramente a estratégia de agrupamento e de organização em unidades decimais desses números. Nos resultados apresentados na subseção 3.1., foi possível observar que os números telefônicos são agrupados e organizados diferentemente. No caso dos números telefônicos de três dígitos, a estratégia recorrente de agrupamento e de organização é do tipo “C”, isto é, tais números são comumente agrupados em uma única unidade entoacional e enunciados como uma unidade centesimal. Assim, por exemplo, o número telefônico “190” é enunciado preferencialmente “cento e noventa”. No caso dos números telefônicos de oito dígitos, a estratégia recorrente de agrupamento é do tipo “UU-UU-UU-UU”, isto é, esses números são frequentemente agrupados em quatro unidades entoacionais, enunciadas em valores unitários. Assim, por exemplo, o número telefônico “34452348” é enunciado preferencialmente “três quatro”-“quatro cinco”-“dois três”-“quatro oito”. Por fim, no caso dos números telefônicos de onze dígitos, a estratégia recorrente de agrupamento é do tipo “UC-UUU-UU-UU”, isto é: quatro unidades entoacionais, sendo a primeira delas enunciada como uma unidade seguida de uma centena, e as demais em unidades. Assim, por exemplo, o número telefônico “08002812112” é enunciado preferencialmente “zero oitocentos”- “dois oito um”- “dois um”- “um dois”. Estes resultados corroboram resultados de estudos semelhantes feitos para outras línguas, no sentido que números telefônicos são em geral organizados em unidades entoacionais relativamente pequenas e são enunciados em unidades decimais de unidades, dezenas e, no máximo, centenas (Amino e Osanai, 2011; Baumann e Trouvain, 2001; Martin, 1997).

Em seguida, procuramos descrever os padrões entoacionais de cada unidade entoacional associada aos três tipos de números telefônicos aqui investigados: números de três, de oito e de onze dígitos. Assim, verificamos que os números de três dígitos, enunciados em apenas uma unidade entoacional, são caracterizados pela combinação de acentos tonais M U D, isto é: um tom médio, seguido de um tom com subida suave, seguido de um tom com descida suave. Este padrão pode ser visualizado no Gráfico 5, feito a partir de valores médios normalizados, com o auxílio do *ProsodyPro*. Note que a configuração obtida corresponde à anotação proposta pelo *INTSINT*.

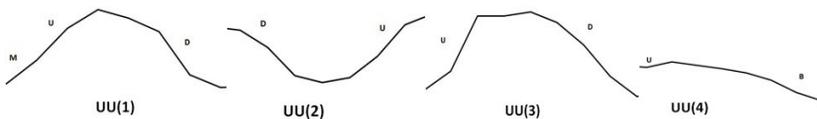
GRÁFICO 5: Contorno entoacional padrão na enunciação dos números de três dígitos.



Os números de oito dígitos, por sua vez, agrupados em quatro unidades entoacionais de dois dígitos, enunciados em unidades, caracterizam-se por padrões entoacionais particulares. A primeira unidade entoacional deste tipo de número telefônico, UU(1), tal como a unidade entoacional de números de três dígitos, é iniciada em um tom médio, seguido de um tom com subida suave, seguido de um tom com descida suave: MUD; a segunda unidade entoacional, UU(2), inicia-se com o mesmo tom com que a unidade entoacional anterior é concluído:

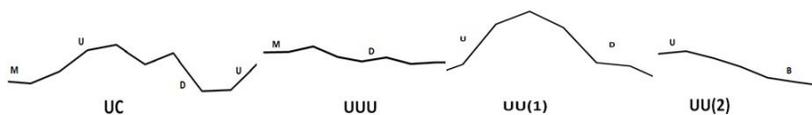
descida suave, que é seguido por um tom com subida suave: DU; a terceira unidade entoacional, UU(3), inicia-se com um tom de subida suave, seguido de um tom com descida suave: UD; por fim, a quarta unidade entoacional, UU(4), é caracterizada por um tom médio seguido de um tom final mais baixo: MB. A configuração do conjunto dessas unidades entoacionais está representada no Gráfico 6:

GRÁFICO 6: Contorno entoacional padrão na enunciação dos números de oito dígitos.



No caso dos números de onze dígitos, a preferência foi um conjunto de quatro unidades entoacionais. A primeira delas, UC, caracteriza-se por um movimento bastante complexo, se comparado com as unidades entoacionais já analisadas: inicia-se com um tom médio, seguido de um tom com subida suave, seguido de um tom com descida suave, finalizando com um tom com subida suave: M U D U; a segunda unidade entoacional, um conjunto de três números enunciados em unidades, UUU, caracteriza-se pela combinação de um tom médio seguido de um tom com subida suave: M D; a penúltima unidade entoacional do agrupamento numérico, UU(1), tem a mesma configuração da curva entoacional da penúltima unidade entoacional dos números de oito dígitos: inicia-se com um tom de subida suave, seguido de um tom com descida suave: U D. Por fim, a quarta e última unidade entoacional, UU(2), inicia-se com um tom de subida suave e conclui com um tom baixo: U B. A configuração do conjunto dessas unidades entoacionais está representada no Gráfico 7:

GRÁFICO 7: Contorno entoacional padrão na enunciação dos números de onze dígitos.



É importante observar que, assim como acontece com agrupamentos de unidades entoacionais no discurso (Couper-Kuhlen, 1986), a enunciação de números telefônicos segue um padrão: unidades entoacionais intermediárias concluem em um tom não-terminal. Apenas a última unidade entoacional (nos casos dos números de oito e onze dígitos) é finalizada com um tom terminal (B).⁸

Conclusões

Os resultados deste trabalho permitiram identificar e descrever contornos entoacionais típicos de números telefônicos comuns no português brasileiro. Foi demonstrado aqui que a entoação apenas parece ser suficiente para caracterizar prosodicamente três tipos de números telefônicos no PB.

Futuramente, seria interessante dar continuidade ao presente estudo e investigar outros parâmetros prosódicos, tais como a duração e intensidade na caracterização dos números telefônicos no PB. Achamos também importante ampliar o *corpus*, aumentar os números de participantes e realizar a coleta de dados em diferentes regiões do Brasil, no intuito de verificar a influência da diversidade regional de português brasileiro sobre a estrutura prosódica dos números telefônicos observada neste estudo. Finalmente, em futuras investigações, planejamos conduzir testes de percepção na base de padrões entoacionais observados nestas

⁸ O mesmo fenômeno foi observado em outras línguas, como, por exemplo, o húngaro (Olaszi, 2000).

análises para testar se os padrões, aplicados à fala sintetizada, são considerados aceitáveis para falantes do PB.

O presente trabalho pode ser considerado uma contribuição para o estudo de enunciação de dígitos numéricos em geral, tópico ainda muito pouco estudado em várias línguas do mundo, e, em particular, para o aprimoramento de sistemas automatizados de reconhecimento e de síntese de fala de números conectados, algo que depende grandemente de informações que este estudo procurou, ainda que parcialmente, sistematizar.

Referências

AMINO, Kanae; OSANAI, Takashi. **Realisation of the prosodic structure of spoken telephone numbers by native and non-native speakers of Japanese**. In proceeding of The 17th International Congress of Phonetic Sciences (ICPhS XVII). Honk kong, China, August p. 17-21, 2011.

BARTKOVA, katarina; JOUVET, Denis. **Selective prosodic post-processing for improving recognition of French telephone numbers**. In Proceedings of THE 6TH EUROSPEECH. Budapest, Hungary, p. 5-9 Sep. 1999.

BAUMANN, Stefan; TROUVAIN, Jürgen. **On the prosody of German telephone numbers**. In proceedings of the 7th conference on speech communication and technology. Aalborg, Denmark, p. 557-560, 2001.

BOERSMA, Paul; WEENINK, David. **Praat: doing phonetics by computer**. Versão 5.3.53. 2013.

CAMPIONE, Estelle ; HIRST, Daniel ; VERONIS, Jean. **Automatic stylisation and symbolic coding of F0: implementations of the INTSINT model**. in A. Botinis (ed.) Intonation. Research and Applications. (Kluwer, Dordrecht). 2000.

CELESTE, Lc. **MOMEL e INTSINT**: uma contribuição à metodologia do estudo prosódico do Português Brasileiro. Dissertação de Mestrado. Belo Horizonte: Faculdade de Letras, Universidade Federal de Minas Gerais; 2007.

CHAFE, W. **Discourse, Consciousness, and Time: The Flow and Displacement of Conscious Experience in Speaking and Writing**. Chicago, University of Chicago Press. 1994.

COUPER-KUHLEN, Elizabeth. **An introduction to English prosody**. Tübingen: Max Niemeyer. 1986.

CRYSTAL, David. **Prosodic Systems and Intonation in English**. Cambridge: The Cambridge University. 1969.

HAKULINEN, Jaakko; TURUNEN, Markku; RAIHA, K-J. **The use of prosodic features to help users extract information from structured elements in spoken dialogue systems**. In Proceedings of the ESCA Tutorial Research Workshop on Dialogue and Prosody. Eindhoven, The Netherlands, Sep.1-3, 1999.

HIRST, Daniel. **A Praat plugin for MOMEL and INTSINT with improved algorithms for modelling and coding intonation**. In Proceedings International Conference on Phonetic Sciences, Saarbrücken, 2007.

HIRST, Daniel; DI CRISTO, Albert. **Intonational Systems, a survey of twenty languages**. Cambridge: Cambridge University Press. 1998.
_____. Modelling French micromelody: analyses and synthesis. *Phonetica*, n. 43, p. 11-30. 1986.

HIRST, Daniel; ESPESSER, Robert. **Automatic Modeling of Fundamental Frequency Using a Quadratic Spline Function**. Aix-Provence: Travaux de l'Institut de Phonétique. 1993.

HIRST, Daniel. **Form and Function in the Representation of Speech Prosody**. Aix-en Provence: Universite de Provence. 2005.

KENT, Raymond; READ, Charles. **The Acoustic Analysis of Speech**. San Diego: Singular Publishing Group Inc. 1992.

LADD, Robert. **Intonational phonology**. Cambridge: Cambridge University Press. 1996.

LIDIANE, Michelle Coelho de Souza. **A Prosódia no Comando Militar**. Dissertação (Mestrado em Linguística). Faculdade de Letras, Universidade Federal de Minas Gerais, 2007.

LOUW, Ja.; BARNARD, Etienne. **Automatic intonation modeling with INTSINT**. Proc. of the 15th Annual Symposium of the Pattern Recognition Association of South Africa. Grabow, November 2004.

MARTIN, Philippe. **La prosodie**. 2007. Disponível em: < <http://www.linguistes.com/phonetique/prosodie.html>>. Acesso em: 20 fev. 2015

NESPOR, Marina; VOGEL, Irene. **Prosodic phonology**. Dordrecht: Foris Publications, 1986.

OLASZI, P. **Analysis of Written and Spoken Form of Hungarian Numbers for TTS Applications**. International Journal of Speech Technology, v. 3, n. 3/4, p. 177–186, 2000.

VAGNER, Latsch; SERGIO Netto. **Obtenção de marcas de pitch em sinais de voz para síntese por concatenação temporal**. In: iii Congresso Brasileiro de Engenharia de Áudio da Aes-brasil apresentado na ix convenção da aes-brasil. São Paulo, p. 11-13 abr. 2005.

VÉRONIS, Jean; HIRST, Daniel; ESPESSER, Robert; IDE, Nancy. **NL and speech in the MULTEXT project**. AAAI '94 Workshop on Integration of Natural Language and Speech., p. 72-78, 1994.

XU, Yi. **ProsodyPro.praat**. (1999-2005).

_____. **ProsodyPro—A Tool for Large-scale Systematic Prosody Analysis**. In Proceedings of tools and Resources for the Analysis of Speech Prosody (TRASP 2013), Aix-en-provence, France. p. 7-10, 2013.

Recebido em 27/03/2015 e Aceito em 10/06/2015.