

PADRÕES TEMPORAIS EM AGRUPAMENTOS DE NÚMEROS TELEFÔNICOS MÓVEIS NO PORTUGUÊS DO BRASIL

TEMPORAL PATTERNS OF MOBILE TELEPHONE NUMBERS IN BRAZILIAN PORTUGUESE

Miguel de Oliveira, Jr¹
Remildo Barbosa da Silva²
Universidade Federal de Alagoas

RESUMO

Este estudo teve por objetivo central descrever padrões temporais da enunciação de números telefônicos móveis no português falado no Brasil, cuja estrutura mais recente caracteriza-se por um agrupamento de nove dígitos. A descrição é baseada na análise acústica dos seguintes fenômenos temporais: ocorrência de pausas, duração de pausas e velocidade de fala. A partir da análise dos dados, foi possível estabelecer um modelo temporal típico de enunciação de números telefônicos móveis de nove dígitos, que é apresentado ao final do trabalho.

PALAVRAS-CHAVE: Prosódia. Enunciado. Número telefônico.

ABSTRACT

This study aimed to describe temporal patterns of the enunciation of mobile telephone numbers in Portuguese spoken in Brazil, whose most recent structure is characterized by a nine-digit grouping. The description is based on the acoustic analysis of the following temporal phenomena: occurrence of pauses, duration of pause and speech rate. From the analysis of the data, it was possible to establish a typical temporal model of enunciation of nine-digit mobile numbers, which is presented at the end of the work.

KEYWORDS: Prosody. Utterance. Telephone number.

INTRODUÇÃO

Há um avanço considerável de estudos sobre as propriedades acústicas de estruturas linguísticas recorrentes, como é o caso de agrupamentos numéricos utilizados como meio de identificação (documentos de identificação pessoal, número de telefone, números em cartões de crédito e números de conta bancária, por exemplo). Isso pode ser verificado pela crescente descrição da estrutura prosódica de números naturais em diversas línguas, tais como o alemão (Baumann & Trouvain, 2001; Creer, 2002), o inglês (Rahim et al., 2001), o espanhol (López, Caminero, Cortázar, & Hernández, 1998), o francês (Mokbel et al., 1997), o japonês (Amino & Osanai, 2011), o húngaro (Olaszi, 2000), o chinês (Wang & Seneff, 1998) e o português europeu (Rodrigues, 2001) e do português do Brasil (Musiliyu, 2015).

A eficiência desses sistemas ainda é, em muitos casos, sofrível. Isso se deve em parte pela escassez de descrições de aspectos linguísticos da fala natural, o que impossibilita que tais sistemas sejam treinados adequadamente. Geralmente, no ato da interação com esses sistemas, é ainda

¹ Professor de Linguística do Programa de Pós-Graduação em Letras e Linguística da Universidade Federal de Alagoas.

² Mestrando em Linguística do Programa de Pós-Graduação em Letras e Linguística da Universidade Federal de Alagoas.

necessário que o falante realize seu enunciado de forma mais lenta que o normal para que haja a decodificação de fala, o que não é o ideal. A síntese de fala pode também não apresentar um bom resultado, por conta da inapropriada segmentação e distribuição de dígitos e da inadequação de aspectos temporais em sua produção: a velocidade de fala e pausas entre dígitos, por exemplo, podem ser considerados pouco naturais.

É importante salientar o impacto que o avanço de estudos que sirvam como base para o aprimoramento de sistemas de síntese e reconhecimento de fala pode ter para a sociedade, principalmente para setores específicos, como o de pessoas com deficiência visual e/ou auditiva, por exemplo. Reconhece-se que sistemas de fala automatizados são de grande valia para esses setores, cujos membros poderão se beneficiar de maneira mais adequada, tendo melhores resultados no processo de transmissão e recepção de informações mediante o uso de sistemas aprimorados.

O objetivo do presente estudo é apresentar uma descrição de características temporais da fala em agrupamentos de números telefônicos móveis de nove dígitos proferidos por falantes do português do Brasil. Especificamente, ocupa-se em descrever a velocidade de fala, a ocorrência e a duração de pausas presentes nesses agrupamentos. O estudo apresenta, como proposta, um modelo de agrupamento numérico para a enunciação de números telefônicos móveis no português brasileiro, a partir de informações temporais³.

1. Estado da arte

Serão apresentados a seguir, de maneira bastante sucinta, alguns conceitos-chave para a discussão do presente estudo. Também um breve painel sobre pesquisas recentes acerca da estrutura e do agrupamento dos números telefônicos será oferecido.

1.1. Prosódia

Segundo Couper-Kuhlen (1986), o termo “prosódia” foi utilizado pela primeira vez pelos gregos para descrever os aspectos da fala que não estão indicados na grafia, especificamente o tom. À medida que os estudos de elementos da fala que não são indicados na grafia foram se aprofundando, o termo “prosódia” passou a ser usado com um significado mais amplo, fazendo com que os estudos prosódicos se estendessem aos demais componentes acústicos da fala: a amplitude e a duração.

De acordo com Oliveira Jr. (2000), o estudo da duração na prosódia está relacionado a qualquer fenômeno de tempo presente na fala, como: o tempo de fonação, a velocidade de articulação, a velocidade de fala, o comprimento de sílabas, a pausa e sua duração. A seguir, serão apresentados e discutidos conceitos duracionais pertinentes ao presente estudo: pausa e velocidade de fala.

1.1.1. Pausa

De acordo com O'Connell & Kowal (1983), a pausa ocorre por vários fatores, tais como: ansiedade, ênfase, interrupções e respiração. As pausas podem ser classificadas em pausas preenchidas e pausas não-preenchidas, levando-se em consideração a ausência ou presença de vocalização em sua produção. No que diz respeito à sua ocorrência como elemento linguístico, ela é usada muito frequentemente na marcação de fronteiras entre unidades de sentido. Desta forma,

³ Este estudo é fruto do projeto “A prosódia de agrupamentos numéricos no português do Brasil, financiado pelo CNPq” (Processo: 405661/2012-6).

de acordo com Oliveira, Jr. (2000), a pausa pode ser entendida como uma pista para a segmentação de enunciados.

Oliveira Jr (2000) observa que geralmente é aceita a ideia de que a pausa reflete a hierarquia sintática da língua: pausas longas ocorrem, na maior parte dos casos, entre frases, do que entre orações. O que muitos estudiosos sugerem é que isso se dá devido a processos cognitivos que são necessários na formulação e na percepção de unidades linguísticas.

Os resultados obtidos por Oliveira Jr (2000) para dados de fala espontânea do português falado no Brasil indicaram a presença de pausas silenciosas em 24% do tempo total de enunciação. Em relação à função da pausa como um elemento acústico de delimitação de constituintes discursivos, ele observou que apenas 45% das unidades (ou frases) entoacionais presentes no seu corpus são seguidas por uma pausa, o que, segundo ele, é um índice relativamente pequeno quando comparado a outros estudos. Para Oliveira Jr (2000), a pausa parece ser um indicador mais expressivo de constituintes hierárquicos maiores, como a unidade discursiva.

1.1.2. Velocidade de fala

A velocidade de fala está intimamente ligada a vários fatores externos. Um fator crucial são as emoções. De acordo com Fairbanks e Hoaglin (1940), a raiva, o desespero, a tristeza e o medo, por exemplo, podem propiciar diferentes níveis de velocidade de fala. Importante notar que isso pode variar de língua para língua.

Essen (1949) observa que os aspectos semânticos também podem influenciar na velocidade de fala e na duração de um enunciado. Ele afirma que o significado tem uma influência decisiva na velocidade de fala, ou seja, ele demonstrou que, quanto mais significado existe num enunciado, a velocidade de articulação tende a ser mais lenta.

O estilo de fala também pode ter um impacto na velocidade de fala. Assim, por exemplo, a narrativa espontânea e a leitura têm diferença considerável de velocidade (Henze, 1953). A idade é outro aspecto que pode apresentar uma variação significativa. Myzak e Hanley (1958) fez estudos com pessoas com diferença de idade bastante significativa. Ele analisou a velocidade de fala de pessoas entre 30 e 90 anos e conclui que a velocidade de fala tende a ser mais lenta em indivíduos mais velhos (ver também: Shipp & Hollien 1972).

Em seu estudo sobre a velocidade de fala em narrativas espontâneas, Oliveira Jr. (2000) observou que, ao contrário do que acontece em várias línguas para as quais o fenômeno já foi sistematicamente estudado, segundo os quais a média de velocidade de fala varia entre 2 a 4 sílabas por segundo, o português falado no Brasil apresenta um valor médio de 5,5 sílabas por segundo, valores relativamente mais altos que os reportados na literatura para outras línguas.

1.2. Números telefônicos no Brasil

No Brasil, os números telefônicos são regulados pela Agência Nacional de Telecomunicações (ANATEL). A representação gráfica “021 82 3522 9679” permite agrupar os números em blocos que possibilitam ser identificados adequadamente. Desta forma, no exemplo anterior os três primeiros números (021) são o código da operadora, os dois dígitos seguintes (82) são o código de área da região, e os últimos números (3522 9679) correspondem ao número do assinante. Neste caso, especificamente, de um número de telefone fixo.

No Brasil, o agrupamento referente à identificação do número do assinante era comumente composto por oito dígitos, mas recentemente as operadoras de telefonia móvel foram obrigadas a acrescentar mais um dígito no início deste agrupamento para suprir as demandas do uso da telefonia móvel.

1.3. Estrutura e agrupamento dos números telefônicos

De acordo com Baumann e Trouvain (2001), os números telefônicos são estruturados em três níveis: agrupamento, expressão e acentuação. O primeiro nível diz respeito à maneira como os números são agrupados. Números telefônicos não são considerados uma única unidade; são constituídos de pequenas unidades, com, no máximo, três elementos por unidade. O segundo nível está relacionado precisamente à forma como esses números serão pronunciados depois de agrupados. Por exemplo: os três primeiros números em 99170 9099 podem ser enunciados como três unidades: “nove, nove, um”, ou como uma unidade e uma dezena: “noventa e nove, um” ou “nove, noventa e um”. O terceiro nível trata exatamente da proeminência na elocução dos números, visto que, de acordo com a teoria da Fonologia Métrica (1996), proeminências em agrupamentos costumam a obedecer um padrão de sequência alternante forte - fraco.

Musiliyu (2015), em estudo sobre características prosódicas de números telefônicos no português brasileiro, apresenta descrições de agrupamentos de números telefônicos de três, oito, e onze dígitos. O autor encontrou várias estratégias de agrupamento para números telefônicos de oito dígitos, sendo uma de grande preferência entre os falantes. Nesse caso, com a preferência de 85% dos casos, os números telefônicos de oito dígitos foram agrupados como quatro agrupamentos binários.

Na análise da distribuição decimal dos números telefônicos, Musiliyu (2015) mostra que a maioria dos falantes prefere utilizar unidades, e não dezenas ou centenas, na enunciação dos números. Desta forma, ele concluiu que os números telefônicos de oito dígitos são agrupados em quatro blocos de dois dígitos, ou seja, de forma binária, sendo enunciados em unidades, independentemente de como os números apareçam dispostos graficamente.

É importante notar que o estudo de Musiliyu (2015), embora seja pioneiro na descrição prosódica de números telefônicos, limita-se a considerações de ordem entoacional, não apresentando nenhuma análise temporal, como é a proposta do presente estudo. Ademais, o estudo não incluiu números de nove dígitos, o que na altura era um dado ainda incomum.

2. Metodologia

O estudo aqui reportado foi realizado a partir de dados de um projeto maior, cujo objetivo foi descrever características prosódicas de vários agrupamentos numéricos do português falado no Brasil. A metodologia descrita a seguir é, em termos globais, a mesma empregada pelo projeto maior, embora foque nas especificidades dos dados que foram aqui utilizados.

2.1. Informantes

Como mencionado acima, os dados utilizados neste estudo fazem parte de um projeto maior, cujo objetivo é estudar características prosódicas de números naturais do português falado no Brasil. O projeto contou, ao todo, com a participação de 120 informantes voluntários, todos falantes nativos do português falado no Brasil. Os participantes tinham diferentes níveis de escolaridade, possuindo minimamente o ensino fundamental incompleto. Ao todo, foram selecionados dados de sessenta e oito informantes, sendo estes 26 homens e 42 mulheres, com idade entre 18 e 50 anos (média 31, desvio padrão 9). É importante observar que dados demográficos foram registrados no estudo somente por uma questão de controle. Não se assumiu variabilidade nos dados de fala coletados baseada em dados demográficos, nem o estudo objetivou lavar dados demográficos em consideração para a análise.

2.2. Coleta de dados

Todos os voluntários participaram de uma entrevista que os estimulava a proferir, naturalmente, vários agrupamentos numéricos, entre os quais documentos de identificação e números telefônicos pessoais. Em seguida, foram instruídos verbalmente a ler, da maneira que julgassem mais adequada, agrupamentos numéricos apresentados em uma série de slides. Esses agrupamentos foram apresentados em um formato padrão, seguidos de uma descrição do tipo de agrupamento (e.g., “Número de telefone”). Os voluntários proferiram todos os agrupamentos duas vezes cada. Foi selecionada a enunciação mais fluente das duas.

A coleta de dados foi feita com um gravador digital de flash Marantz PMD661, utilizando-se um microfone tipo headset. Os arquivos de áudio estão em formato PCM, não-comprimido (gravados em formato WAV), com taxa de amostragem de 96 kHz e 32 bits por amostra.

2.3. Estímulo

Os agrupamentos numéricos apresentados como estímulo para leitura eram agrupamentos numéricos reais. Os números telefônicos, em particular, encontravam-se, na época da coleta de dados, ativos. Para este estudo, foram selecionados os dois únicos números telefônicos de nove dígitos que foram apresentados aos informantes como estímulo no projeto maior: 98543-6721 e 99255-2345. Decidimos utilizar neste estudo ambos os números, para que assim fosse possível observar se uma mesma estratégia de agrupamento é utilizada em ambos os casos, o que nos permite propor uma generalização mais adequada.

2.4. Unidades de medida

A unidade de medida mais comum em estudos de velocidade da fala é a sílaba por segundo. Apesar de apresentar alguns inconvenientes, uma vez que comumente o falante exclui sílabas ou as incorpora a outra palavra de um enunciado, esta medida é descrita pela literatura como a menos problemática. No que diz respeito à pausa, há uma grande discussão acerca de qual será a duração mínima para que um período de silêncio em fala fluente seja considerado pausa. Glukhov (1975), por exemplo, a partir de um estudo feito em diversas línguas, mostrou que existe uma grande diferença de valores de pausa nas mais diversas línguas, e que esses valores variam entre 50 a 150 ms. Para ele, não é possível estabelecer um valor mínimo universal para pausa. Neste estudo, a pausa foi considerada como um período de silêncio igual ou superior a 150 ms.

2.5. Anotação dos dados

Os dados coletados foram anotados e analisados no aplicativo computacional Praat (BOERSMA; WEENINK, 2017). A anotação permitiu que as informações dos dados coletados fossem devidamente anotados em diferentes camadas. A primeira camada corresponde à transcrição ortográfica da cadeia numérica enunciada e está segmentada em palavras. A segunda camada informa a quantidade de dígitos proferidos pelo falante, por agrupamento. Cada agrupamento nesta camada corresponde a uma unidade entoacional⁴. A camada seguinte, também segmentada em unidades entoacionais, traz informações da distribuição decimal de cada agrupamento: unidade (U), dezena (D) ou centena (C). A quarta camada traz informação do número de sílabas em cada agrupamento. É importante ressaltar que a quantidade de sílabas registradas nesta camada corresponde a sílabas de caráter fonético e não fonológico ou gráfico. A

⁴ 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. (CHAFFE, 1994).

quinta camada traz informações acerca do tipo de agrupamento numérico (telefone celular, CEL, por exemplo), e da posição em que se encontra dentro da cadeia. A última camada foi inserida automaticamente pelo Praat para indicar a existência (S) ou inexistência (N) de pausas no arquivo de áudio.

3. Resultados e discussões

A seguir serão apresentados os resultados de análises feitas para propriedades de agrupamento, pausa e velocidade de fala obtidos. Os resultados serão discutidos em conjunto, logo a seguir.

3.1. A estratégia de agrupamento

O gráfico 1 abaixo traz informações acerca das estratégias mais comuns, baseada em 136 enunciados.

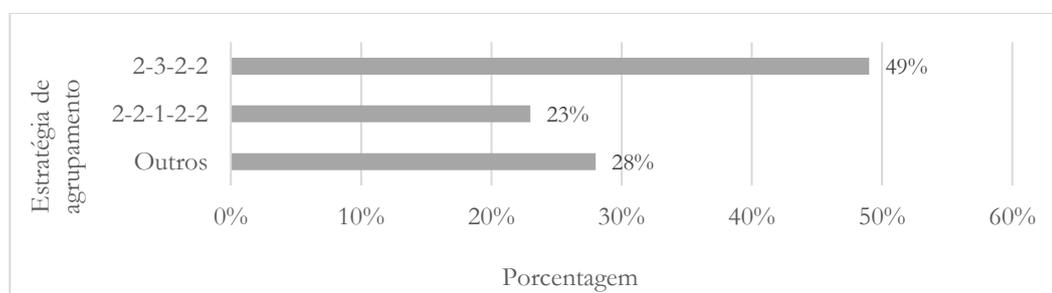


Gráfico 1. Resultado da análise da estratégia de agrupamento para números telefônicos

Conforme exposto no gráfico 1, a estratégia de agrupamento de números telefônicos de nove dígitos mais comum é a do tipo 2-3-2-2: quatro unidades entoacionais, na qual a primeira é constituída por um grupo binário, a segunda por um grupo ternário e as duas últimas por um grupo binário cada. Com 23% dos casos, a estratégia de agrupamento 2-2-1-2-2 ocupou o segundo lugar entre estratégias mais comuns. Os outros dez tipos de agrupamentos encontrados somaram 28% de ocorrência. O teste estatístico do Qui-quadrado (χ^2) indicou que as duas estratégias de agrupamento mais recorrentes são estatisticamente diferentes ($\chi^2 = 13.2245$, $p = 0.0003$, $df = 1$). Importante notar que a estratégia de agrupamento 2-3-2-2 ocorreu com maior frequência também em se considerando os números individualmente. Para o número 98453-6721, a estratégia foi empregada em 54% dos casos. Já para o número 99255-2345, ocorreu em 44% dos casos.

3.2. A distribuição decimal

A distribuição decimal foi analisada na estratégia de agrupamento mais significativa, 2-3-2-2, que considerou os 67 enunciados resultantes. O gráfico 2 abaixo mostra os resultados encontrados:

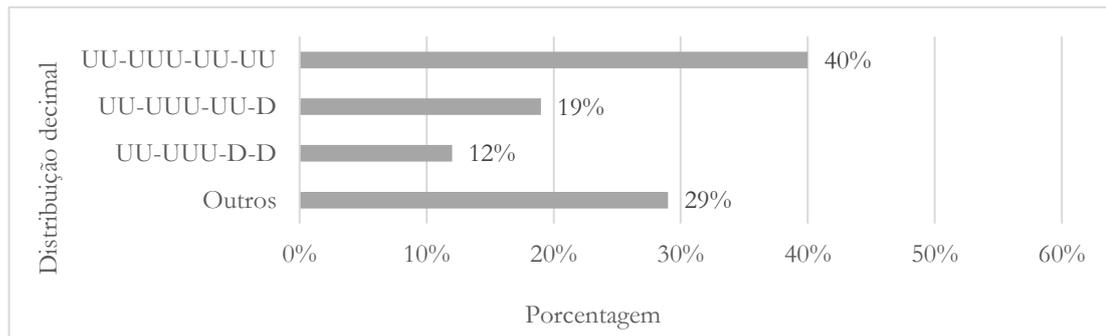


Gráfico 2. Resultado da análise da estratégia de distribuição decimal

Como pode ser observado no gráfico acima, os resultados apontaram para uma maior frequência de uso de unidades (U) como estratégia de distribuição decimal para a estratégia de agrupamento 2-3-2-2. A distribuição decimal mais frequente é caracterizada pela ocorrência exclusiva de unidades: UU-UUU-UU-UU. As duas outras distribuições decimais que apareceram com uma frequência maior que 10% dos casos incluem dezenas. O teste estatístico do Qui-quadrado (χ^2) indicou que as duas distribuições decimais mais recorrentes são estatisticamente diferentes ($\chi^2 = 4.9000$, $p = 0.0269$, $df = 1$).

3.3. A ocorrência e duração de pausa

A análise para a ocorrência e duração de pausa foi feita com base nos 67 enunciados agrupados em 2-3-2-2. Foi possível verificar que pausas ocorrem em 63% desses dados. Desse total, há ocorrência de apenas uma pausa por número telefônico em 71% dos casos. Essa pausa ocorre em geral depois da segunda unidade entoacional, UE2 (80% dos casos). Nos demais casos, ocorre depois da primeira unidade entoacional, UE1. Quando há duas pausas, a primeira sempre ocorre depois da UE1, e a segunda, em geral, depois da UE2 (83% dos casos).

Se considerarmos exclusivamente os casos em que há apenas uma pausa no enunciado, observamos que quando a pausa ocorre logo após a UE2, ela costuma ser, em média, maior que quando esta ocorre após a UE1. O gráfico abaixo ilustra o resultado⁵:

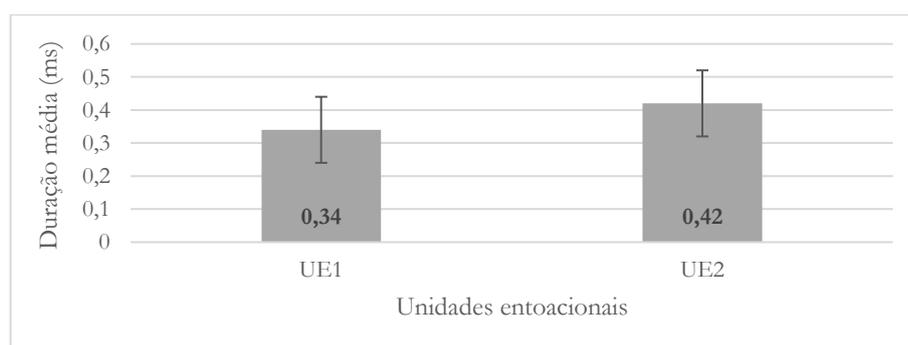


Gráfico 3. Duração média, em milissegundos, das pausas após unidades entoacionais 1 (UE1) e 2 (UE2) nos números telefônicos com agrupamento 2-3-2-2

É importante observar que essa diferença de valores de média reportada no Gráfico 3 acima, embora demonstre uma tendência, não chega a ser estatisticamente significativa ($t = 0.726115$; $df = 28$; $p = 0.4738$).

⁵ Este gráfico, como todos os outros que seguem (de 3 a 10) apresentam barra de erro com valores de erro padrão.

Também foram realizadas análises especificamente para números telefônicos com agrupamento 2-3-2-2 e com distribuição decimal UU-UUU-UU-UU. Há ocorrência bastante expressiva de pausas em números telefônicos enunciados com distribuição decimal UU-UUU-UU-UU, no qual 74% deles apresentaram pausa entre unidades entoacionais.

Seguiu-se a mesma tendência de frequência de pausas para distribuição decimal “UU-UUU-UU-UU”, como era de se esperar. Na maioria dos casos (75%), verificou-se a ocorrência de apenas uma pausa na enunciação do número telefônico. Nesses casos, a pausa ocorreu mais frequentemente (em 56% das realizações) após a UE2. As demais ocorreram após a UE1.

Nos casos de ocorrência de duas pausas na distribuição decimal UU-UUU-UU-UU, a primeira sempre ocorre depois da UE1, e a segunda, em geral, depois da UE2 (83% dos casos).

Quando ocorre isoladamente, a pausa na distribuição decimal UU-UUU-UU-UU não apresenta duração média distinta de acordo com a sua localização. O Gráfico 4 abaixo traz valores médios de pausas que ocorrem depois da UE1 e da UE2. A diferença entre esses valores não é estatisticamente significativa ($t = 0.127537$; $df = 13$; $p = 0.9005$).

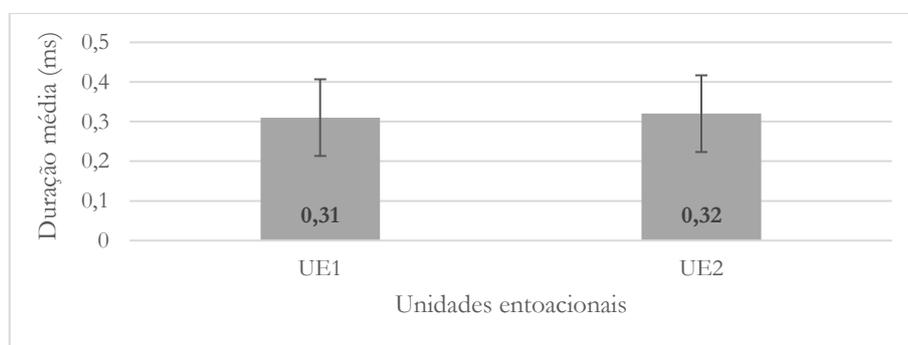


Gráfico 4. Duração média, em milissegundos, das pausas após unidades entoacionais 1 (UE1) e 2 (UE2) nos números telefônicos com distribuição decimal UU-UUU-UU-UU

Um dado curioso pode ser observado no gráfico 5 abaixo: a pausa que ocorre depois da UE1 quando é a única pausa no enunciado é, em média, maior, se comparada à pausa que ocorre na mesma posição em enunciados com duas pausas. A diferença, no entanto, não é estatisticamente significativa ($t = -1.34103$; $df = 16$; $p = 0.1986$).

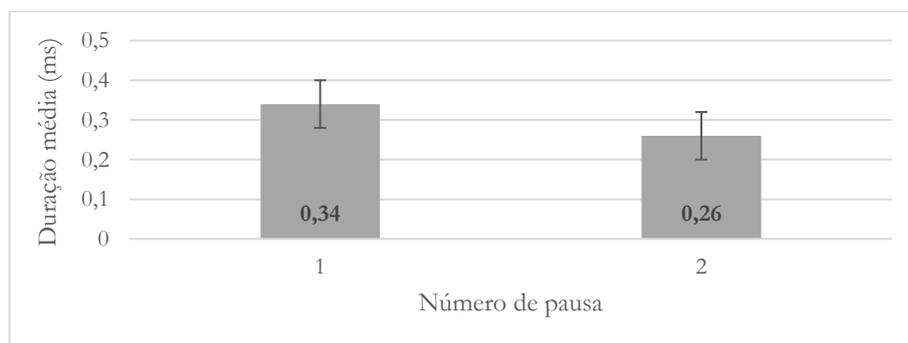


Gráfico 5. Duração média, em milissegundos, das pausas após a unidade entoacional UE1 nos números telefônicos com agrupamento 2-3-2-2 que apresentam uma pausa ou duas pausas

A pausa que ocorre depois da UE2, por outro lado, costuma ser, em média, maior em enunciados com duas pausas, como mostra o gráfico 6 abaixo. As diferenças entre as médias nesses casos não são estatisticamente significativas ($t = 1.3058968$; $df = 32$; $p = 0.2975$), mas os valores de duração mostram uma tendência neste sentido.

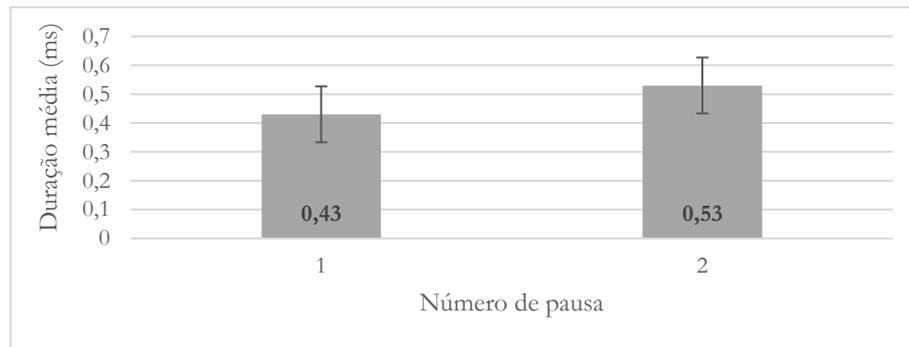


Gráfico 6. Duração média, em milissegundos, das pausas após a unidade entoacional UE2 nos números telefônicos com agrupamento 2-3-2-2 que apresentam uma pausa ou duas pausas

Nos casos em que duas pausas ocorrem no mesmo enunciado, a pausa que ocorre depois da UE2 é, em média, maior que as pausas que ocorrem depois de UE1, como ilustra o gráfico 7 abaixo. De acordo com um teste t para amostras independentes, a diferença entre as médias de valores de pausa, levando-se em conta a sua localização, é estatisticamente significativa: $t = 3.661091$; $df = 20$; $p = 0.0016$.

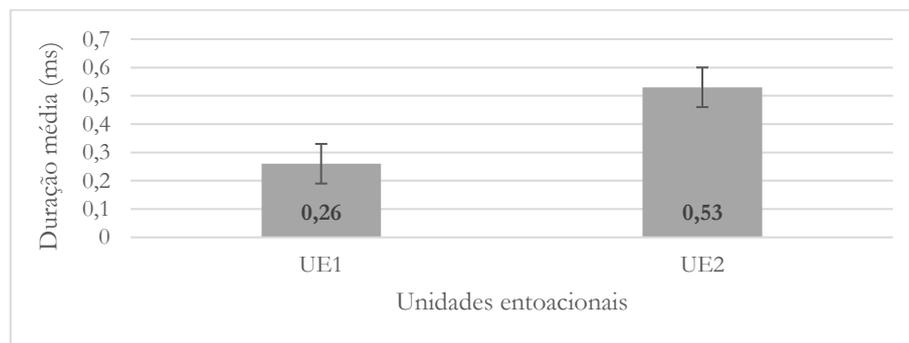


Gráfico 7. Duração média, em milissegundos, das pausas após UE1 e UE2 ocorridas num mesmo enunciado de números telefônicos com agrupamento 2-3-2-2

Filtrados os resultados apenas para a distribuição decimal UU-UUU-UU-UU, é possível observar a mesma tendência: pausas são, em média, mais longas após UE2 em enunciados com duas pausas, como ilustra o gráfico 8 abaixo.

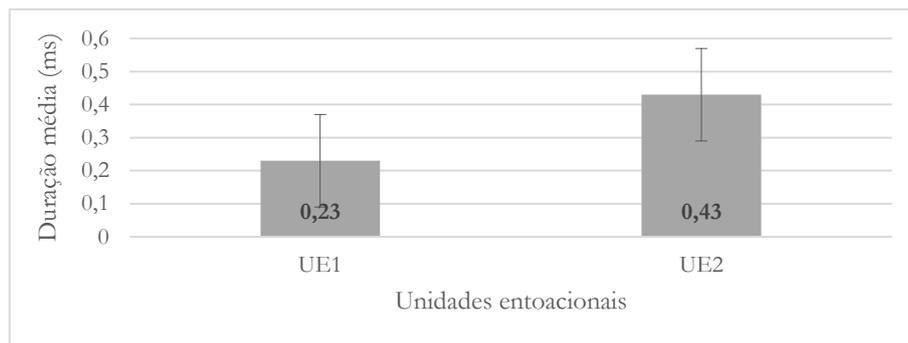


Gráfico 8. Duração média, em milissegundos, das pausas após UE1 e UE2 ocorridas num mesmo enunciado de números telefônicos com distribuição decimal UU-UUU-UU-UU

No entanto, de acordo com um teste *t* para amostras independentes, a diferença entre as médias de valores de pausa, levando-se em conta a sua localização, não é, neste caso, estatisticamente significativa para a distribuição decimal UU-UUU-UU-UU: $t = 1.470991$; $df = 8$; $p = 0.1795$. Este resultado está provavelmente relacionado à pequena quantidade de dados restantes, quando filtrados para que se considerem apenas a distribuição decimal UU-UUU-UU-UU, e pela alta variância de valores nesta amostra restrita (refletida na barra de erro do Gráfico 8).

3.4. A velocidade de fala dos números telefônicos de nove dígitos

A análise da velocidade de fala nas unidades entoacionais dos números telefônicos foi feita apenas para o agrupamento que ocorreu com maior frequência nos dados coletados: 2-3-2-2. Nestes, análise específica da distribuição decimal mais recorrente (UU-UUU-UU-UU) também foi realizada⁶.

O gráfico abaixo mostra os resultados para os 67 enunciados em análise:

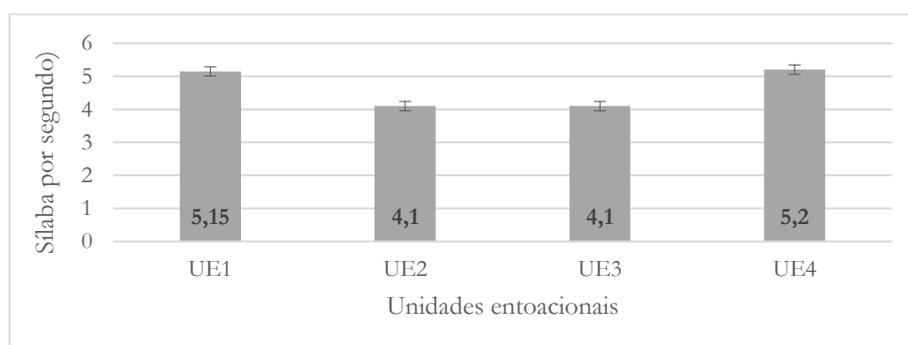


Gráfico 9. Velocidade de fala, em sílaba por segundo, nas unidades entoacionais do agrupamento 2-3-2-2

É possível observar no gráfico 9 acima que a velocidade de fala varia de UE para UE. A tendência é ser maior nas UE1 (5.15 s/s) e UE4 (5.2 s/s) e menor nas UE2 (4.1 s/s) e UE3 (4.1 s/s). Houve uma diferença estatisticamente significativa entre os grupos entoacionais, conforme determinado pelo teste ANOVA de uma via: $F(3,264) = 21.6035$, $p < 0.001$. Testes de pos-hoc Tukey-Kramer HSD revelaram que UE2 e UE3 diferem significativamente de UE1 e UE4, a um

⁶ É importante observar que os dados para esta análise foram padronizados com valores z-scores, mas os resultados foram os mesmos se comparados com valores não-padronizados.

nível de $p < 0.001$. No entanto, nem as duas UEs intermediárias, nem as duas UEs nos extremos diferem entre si.

O gráfico abaixo mostra os resultados para a distribuição decimal classificada como UU-UUU-UU-UU:

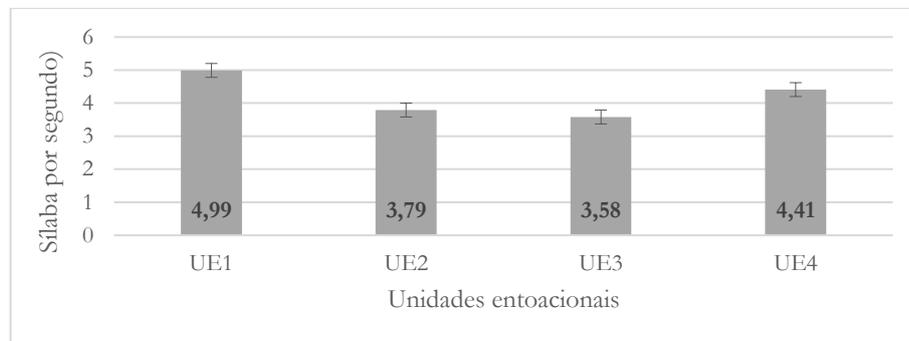


Gráfico 10. Velocidade de fala, em sílabas por segundo nas unidades entoacionais da distribuição decimal UU-UUU-UU-UU

Os resultados para a distribuição decimal UU-UUU-UU-UU seguem a mesma tendência daquilo que foi observado no agrupamento 2-3-2-2 em geral: unidades entoacionais das extremidades são caracterizadas por velocidade de fala maior que unidades entoacionais intermediárias. Houve aqui também uma diferença estatisticamente significativa entre os grupos entoacionais, conforme determinado pelo teste ANOVA de uma via: $F(3,104) = 8.9537, p < 0.001$. Testes de pos-hoc Tukey-Kramer HSD revelaram que UE2 e UE3 diferem significativamente de UE1 e UE4, a um nível de $p < 0.001$. Aqui também, não foram encontradas diferenças entre as duas UEs intermediárias, nem entre as duas UEs nos extremos.

3.5. Discussões

A despeito de existirem várias estratégias de agrupamento para a enunciação desses números, uma delas foi mais recorrente nos dados analisados: o agrupamento em quatro unidades entoacionais, sendo a primeira delas com padrão binário, a segunda com padrão ternário e as duas últimas com padrão binário: 2-3-2-2. Os resultados corroboram com os reportados por Musiliyu (2015), mostrando que, mesmo com a inclusão de um novo dígito, a estratégia de agrupamento em quatro unidades entoacionais permanece como a mais recorrente.

No que diz respeito à distribuição decimal no agrupamento do tipo 2-3-2-2, observou-se a utilização exclusiva de unidades nos enunciados como estratégia mais recorrente: UU-UUU-UU-UU. Esta estratégia de enunciação também corrobora com o que foi encontrado por Musiliyu (2015): todos os números da cadeia numérica de maior preferência foram enunciados em forma de unidades. Os resultados, como um todo, vão ao encontro daquilo que afirma Baumann e Trouvain (2001): números telefônicos não são enunciados como um todo absoluto, apresentando, em geral, possibilidades de agrupamento que são mais recorrentes e que estão relacionadas à tipologia rítmica da língua.

No que concerne à ocorrência de pausas no agrupamento 2-3-2-2, os resultados mostraram que quando há apenas uma pausa, esta geralmente ocorre após a segunda unidade entoacional (80% dos casos). Importante observar que na representação gráfica do número telefônico há, em geral, uma barra simples (-) delimitando os cinco primeiros números da cadeia numérica dos quatro últimos números. Este ponto corresponde precisamente ao lugar onde, na enunciação, ocorre pausa: entre a UE2 e a UE3.

Em relação a enunciados com duas pausas, a primeira sempre ocorre após a UE1. Uma possível explicação para isto pode ser o fato de a UE2 ser formada por um grupo ternário e não por um grupo binário, como acontece para números de oito dígitos. Isso pode ter causado estranhamento para quem ainda não está habituado a números de nove dígitos⁷, o que fez com que o falante hesitasse, momentaneamente, refletindo assim uma dificuldade momentânea de processamento de informação (GOLDMAN-EISLER, 1968). A segunda pausa, assim como enunciados com apenas uma pausa, também ocorre após UE2, em 83% dos casos.

Nos enunciados com duas pausas, a duração média da pausa após a UE1 foi de 0.26 ms, enquanto que a pausa após a UE2 teve uma duração média de 0.53 ms. Essa diferença é estatisticamente significativa. Alguns estudos sugerem que pausas mais longas são empregadas para dar mais tempo para o falante formular o enunciado seguinte, e, ao mesmo tempo, para ajudar o ouvinte a compreender melhor o conteúdo do enunciado anterior (Ford 1982; Goldman-Eisler, 1961; Miller, Galanter e Pribram 1960; Siegman 1979). Como geralmente quem enuncia um número telefônico o faz para que o ouvinte registre esse número, é possível supor que a pausa mais longa após a UE2 se justifique pragmaticamente, uma vez que ocorre aproximadamente a meio do enunciado, o que permitiria o ouvinte processar mais adequadamente a informação contida em ambas as partes do mesmo.

Em relação à ocorrência de pausa especificamente na distribuição decimal UU-UUU-UU-UU, observou-se que a tendência se manteve. Houve, porém, um ligeiro aumento na ocorrência de pausas. Dos 27 enunciados considerados, 74% apresentou pausa; desses, 75% apresentou apenas uma pausa. Em UU-UUU-UU-UU, a pausa ocorreu com maior frequência após a UE2 (56% dos casos). Aqui também o padrão duracional se repete, a pausa após a UE2 teve média de duração maior que a pausa após a UE1: 0.43 ms e 0.23 ms, respectivamente. A diferença de duração aqui, no entanto, não é estatisticamente significativa, o que pode estar relacionado ao pequeno dado amostral resultante.

Um dado curioso observado nos resultados de duração das pausas foi o fato de a média da duração da pausa variar de acordo com o número de pausas existentes no enunciado. Assim, por exemplo, a pausa que ocorre depois da UE1 quando é a única pausa no enunciado costuma ser, em média, mais longa, se comparada à pausa que ocorre na mesma posição em enunciados com duas pausas. A pausa que ocorre depois da UE2, no entanto, costuma ser, em média, mais longa em enunciados com duas pausas. Essas diferenças, embora não sejam estatisticamente significativas, podem estar relacionadas a uma estratégia compensatória: as pausas que ocorrem após a UE1 em enunciados com duas pausas são menos longas porque as pausas que ocorrem após a UE2 nesse mesmo contexto são mais longas.

Resultados de análise da velocidade de fala nos enunciados revelaram que as unidades entoacionais localizadas aos extremos dos enunciados apresentaram uma média de velocidade de fala mais rápida tanto para 2-3-2-2 quanto para UU-UUU-UU-UU. A velocidade de fala mais lenta, em média, ocorreu nas unidades entoacionais intermediárias. Não há diferença significativa entre a velocidade de fala associada às unidades entoacionais nas extremidades dos agrupamentos numéricos (UE1 e UE4); O mesmo acontece nas unidades intermediárias.

Há possíveis explicações para a velocidade de fala ser maior nas UE1 e UE4. Uma delas está relacionada à ativação de conhecimento prévio. Os números das operadoras de telefonia móvel podem ser reconhecidos pela banda⁸ do número de discagem. O número 9, que foi recentemente incluído no início dos números de telefone móvel, corresponde a um número de banda muito comum no nordeste do Brasil, especificamente em Alagoas, local de coleta dos dados. Estudos têm mostrado que a velocidade de fala está relacionada a processamentos cognitivos no acesso de

⁷ A coleta de dados foi realizada antes da implantação do número de nove dígitos na região em que foi realizada. Números de nove dígitos, no entanto, já eram comuns em outras regiões do país.

⁸ Termo que especifica a faixa de frequências de um determinado sinal antes que ele sofra qualquer tipo de modulação. Em se tratando de números telefônicos, são os dois primeiros números da cadeia numérica.

informações novas ou conhecidas: enunciados mais rápidos estão em geral associados à enunciação de informações conhecidas (Henderson, Goldman-Eisler & Skarbek 1966, Goldman-Eisler 1967, Butterworth & Goldman-Eisler 1979). Além disso, é importante considerar que, tal como afirmado acima, o agrupamento de números telefônicos de oito dígitos ocorre por grupos binários. Esse padrão, que vem sendo usado desde a segunda metade da década de 1990, pode ser responsável por provocar uma velocidade de fala mais rápida no início do enunciado, uma vez que é imediatamente reconhecido pelo falante/leitor. Isso também explicaria a velocidade mais lenta na enunciação da UE2: o agrupamento ternário aí é relativamente novo, o que pode demandar maior esforço cognitivo, resultando em uma enunciação mais lenta. No que diz respeito à também mais lenta velocidade de fala na UE3, é possível que ela se justifique, assim como a ocorrência de pausa imediatamente antes de sua enunciação, pela presença da barra (-), que separa os cinco primeiros números dos quatro últimos na disposição gráfica desses números telefônicos. Essa barra também pode demandar um maior esforço cognitivo no ato da leitura, o que determinaria uma enunciação mais lenta da unidade entoacional seguinte. Por fim, a última unidade entoacional não está associada a nenhum elemento gráfico especial, sendo caracterizada por um agrupamento binário típico, o que justificaria a sua enunciação mais rápida.

Importante observar que as unidades entoacionais UE1 e UE4 apresentaram valores de velocidade de fala que se aproximam dos resultados encontrados por Oliveira Jr. (2000) em um estudo com narrativas espontâneas. Em seus resultados, o autor encontrou uma média de 5.5 sílabas por segundo para média de velocidade de fala no português do Brasil, média alcançada pelas unidades entoacionais UE1 (5.15 s/s) e UE4 (5.2 s/s) no agrupamento 2-3-2-2.

De acordo com Henze (1953), o estilo de fala pode determinar a velocidade a ele associada. O fato de os enunciados aqui analisados apresentarem uma média de velocidade de fala muito semelhante àquela encontrada por Oliveira Jr. (2000), a despeito de serem caracterizados por fala lida e não por fala espontânea, pode ser justificado pelo conteúdo relativamente curto e predizível a eles associados. Essen (1949) afirma que quanto mais significado existe num enunciado, a velocidade de fala tende a ser mais lenta. Como os números telefônicos em si não estão associados a significados complexos, parece ser plausível que sejam, portanto, enunciados em uma velocidade mais rápida, típica de fala espontânea, em oposição à fala lida.

CONCLUSÃO

Este estudo apresenta resultados de uma análise baseada na enunciação de dois exemplares de números telefônicos que seguem o padrão atual para números de telefonia móvel. As análises foram realizadas a partir da enunciação desse par de números por 68 falantes da variedade alagoana do português falado no Brasil, num momento em que números de nove dígitos estavam começando a ser implementados na região. A partir das análises organizacionais e prosódicas realizadas, observou-se um padrão de agrupamento e distribuição decimal para números telefônicos analisados. Os resultados das análises organizacionais apontaram para uma ocorrência mais frequente de agrupamento no formato 2-3-2-2, caracterizado por uma unidade binária, seguido de uma unidade ternária e por mais duas unidades binárias. A distribuição decimal dos números se deu preferencialmente em unidades: UU-UUU-UU-UU, como em “nove oito – cinco quatro três – seis sete – dois um”, por exemplo.

As análises prosódicas aqui realizadas se concentraram em fenômenos temporais, especificamente na ocorrência e duração de pausas, e na velocidade de fala dos enunciados. Os números telefônicos foram enunciados com ou sem pausa; a presença de pausas entre suas unidades constituintes, no entanto, mostrou-se bastante expressiva. O número máximo de pausa encontrado na enunciação de um número telefônico de nove dígitos foi duas. Porém, enunciados com apenas uma pausa foram mais frequentes. A ocorrência de pausa mostrou-se mais significativa na distribuição decimal UU-UUU-UU-UU, sendo sua ocorrência mais expressiva que no

agrupamento 2-3-2-2, a que está vinculada.

O local de ocorrência mais frequente da pausa foi após a UE2, mesmo local onde é localizado o sinal visual demarcatório na versão gráfica dos números telefônicos. Quando uma pausa ocorre no enunciado, ela é em geral localizada após a UE2. Quando duas pausas ocorrem, elas são localizadas após a UE1 e UE2. A duração da pausa em enunciados com apenas uma pausa pode variar, de acordo com a sua localização, sendo maior quando ocorre após a UE2. Em enunciados com duas pausas, a segunda, que em geral ocorre depois da UE2, costuma ter maior duração que a primeira.

A velocidade de fala apresenta variação de UE para UE. A unidade entoacional que foge do padrão binário e a presença de sinal gráfico não-numérico podem ter induzido uma desaceleração da velocidade de fala típica do português falado no Brasil, tal como reportado em estudo com fala espontânea (Oliveira Jr, 2000). Velocidades de fala típicos ocorrem nas unidades entoacionais situadas nos extremos do agrupamento típico de números telefônicos, enquanto que velocidades mais lentas ocorreram em unidades entoacionais intermediárias. Essa tendência foi observada tanto no agrupamento 2-3-2-2 quanto, em um subgrupo deste tipo de agrupamento: aquele cuja distribuição decimal era do tipo UU-UUU-UU-UU.

Os achados aqui reportados, embora estejam relacionados a um conjunto de itens experimentais bastante restrito, corroboram descrições já feitas para agrupamentos numéricos do português falado no Brasil (MUSILYU, 2015). Como observado, não houve variação significativa nos resultados obtidos para os itens experimentais, o que aponta para a generalização de observações. Idealmente, a análise deve incluir vários outros exemplares e, além disso, exemplares de números enunciados espontaneamente. Entretanto, apresentaremos aqui, a partir dos resultados deste estudo seminal, um modelo de enunciação estrutural e temporal de números telefônicos de nove dígitos para o português do Brasil:



Figura 1. Modelo de agrupamento e enunciação temporal de números telefônicos de nove dígitos.

Acima, a figura 1 representa uma proposta de estrutura padrão para agrupamentos de números telefônicos de nove dígitos, no qual estão expressos valores temporais médios de velocidade de fala e de pausa. O modelo proposto pode ser útil, por exemplo, para a implementação em sistemas de síntese de fala. Como investigação futura, propõe-se a realização de testes de aceitabilidade realizados a partir da síntese de exemplares de agrupamentos numéricos baseada no modelo acima proposto.

Este estudo teve como objetivo analisar a estratégia de agrupamento e distribuição decimal e características temporais da fala no novo formato de números de telefone móvel no português do Brasil. A partir dos achados, um modelo de enunciação temporal de números telefônicos com esta característica foi proposto. Configura-se, neste sentido, como uma contribuição na descrição de características prosódicas de estruturas linguísticas recorrentes no português falado no Brasil, uma área de estudo ainda incipiente.

REFERÊNCIAS

- AMINO, K.; OSANAI, T. **Realisation of the prosodic structure of spoken telephone numbers by native and non-native speakers of Japanese** (pp. 236–239). Presented at the ICPHS XVII, Hong Kong, 2011.
- ANATEL. <http://www.anatel.gov.br/setorregulado/index.php?option=com_content&view=article&id=321&Itemid=409>. Acesso em 16/07/2015.
- BARBOSA, P. A. **Prosódia: uma entrevista com Plínio A. Barbosa**. *ReVEL*, v. 8, n. 15, 2010. Disponível em: <http://www.revel.inf.br/files/entrevistas/revel_15_entrevista_plinio.pdf>. Acesso: 20/03/2016.
- BAUMANN, S.; TROUVAIN, J. **On the prosody of German telephone numbers**. In Proceedings of THE 7TH CONFERENCE ON SPEECH COMMUNICATION AND TECHNOLOGY. Aalborg, Denmark, 2001. P. 557-560. Disponível em: <<http://www.coli.uni-saarland.de/publikationen/softcopies/Baumann:2001:PGT.pdf>>. Acesso: 20/03/2016.
- BENUS, S. et al. **Pauses in deceptive speech**. Proceedings of the 3rd Speech Prosody Conference, 2006.
- BOERSMA, P. & WEENINK, D. Praat: doing phonetics by computer [Aplicativo computacional]. Disponível em: <http://www.praat.org/>. 2017.
- COUPER-KUHLEN, E. **An introduction to English prosody**. Tübingen: Max Niemeyer, 1986.
- CRUTTENDEN, A. Preliminaries. In: Alan Cruttenden. **Intonation**. Cambridge University Press, 1986.
- DELFINO, A. **Estudos prosódicos das disfluências de reparo**. Dissertação [Mestrado em Linguística] 129 pág. Universidade Federal de Minas Gerais, 2009.
- FORD, M. Sentence planning units: Implications for the speaker's representation of meaningful relations underlying sentences. In J. Bresnan. **The Mental Representation of Grammatical Relations**. Cambridge, Mass., MIT Press, 1982.
- GOLDMAN-EISLER, F. **The distribution of pause duration in speech**. *Language and Speech* 4: 232-237, 1961.
- GOLDMAN-EISLER, F. **Psycholinguistics: experiments in spontaneous speech**. London, New York: Academic Press, 1968.
- LADD, D.R., *Intonational Phonology*. CUP, 1996.
- LÓPEZ, E. *et al.* **Improvement on connected numbers recognition using prosodic information** (Vol. 2). Presented at the 5th International Conference on Spoken Language Processing, Prosody and Emotion, 1998.
- MILLER, G. A., GALANTER, E. & PRIBRAM, K. H. **Plans and the Structure of Behavior**. New York, Holt, Rinehart and Wiston Inc, 1960.

MOKBEL, C. *et al.* **Towards improving ASR robustness for PSN and GSM telephone applications.** *Speech Communication*, 23(1-2), 141–159, 1997.

MUSILYU, O; OLIVEIRA JR, M. **Padrões entoacionais dos números telefônicos no português brasileiro.** *Revista da ABRALIN*, v.14, n.1, p. 497-532, jan./jun. 2015.

OLIVEIRA JR., M. **Prosodic features in spontaneous narratives.** Tese [Doutorado em Linguística] 275 p. Department of Linguistics - Simon Fraser University, Canadá, 2000.

OLIVEIRA JR., M. *et al.* **A prosódia de agrupamentos numéricos no português do Brasil.** Processo 405661/2012-6. 2012 – 2015.

RAHIM, M. *et al.* **Robust numeric recognition in spoken language dialogue.** *Speech Communication*, 34 (1), 195–212, 2001.

RODRIGUES, F. S. P. **Reconhecimento Robusto de Dígitos e Números Naturais.** Tese [doutorado]. Universidade Técnica de Lisboa, 2001.

SIEGMAN, A. W. Cognition and hesitation in speech. In A. Siegman & S. Felstein. **Of Time and Speech.** Hillsdale, NJ, Lawrence Erlbaum: 151-178, 1979.

WANG, C.; SENEFF, S. **A study of tones and tempo in continuous Mandarin digit strings and their application in telephone quality speech recognition.** *Proc ICSLP'98*, 1998.

Recebido em 19/08/2017

Aceito em 27/09/2017

Publicado em 05/10/2017