Intonational Patterns of Telephone Numbers in Brazilian Portuguese

Oyedeji Musiliyu [1], *Miguel Oliveira Jr.* [2]

[1, 2] Universidade Federal de Alagoas, Alagoas, Brasil

bodeses@yahoo.fr, miguel@fale.ufal.br

Abstract

The main purpose of this paper was to identify intonational patterns of a quite common type of numeric grouping in Brazilian Portuguese: the one associated with telephone numbers. To this aim, 30 samples of spoken telephone numbers, read aloud by 85 native speakers of Brazilian Portuguese were analysed. The description of the intonation contours was observed by using *Momel/Intsint* [1] and *ProsodyPro* [2] scripts for *Praat* (version 5.3.53) [3], through a semi-automatic analysis of pitch variations in numeric groupings that form the telephone numbers. The results show a pattern of intonation and numeric grouping strategy that are sufficient enough to characterize prosodically different types of spoken telephone numbers in Brazilian Portuguese.

Keywords: telephone numbers, intonation, Brazilian Portuguese.

10.21437/SpeechProsody.2014-156

1. Introduction

In the past few years, speech technology advancements have made frequent the use of automatic speech recognition and synthesis systems in multiple applications. Some services based on these automated systems make use of concatenated numeric digits for various purposes, such as: activation of credit cards, banking information, telephone directories, booking inquiries and assistance services for blind and visually impaired individuals, for example.

Services that make use of numerical groupings as input or output data depend on a good voice or textual data processing system of alphanumeric digits information. An efficient system will recognize adequately spontaneous speech and produce a voice or textual output corresponding to adequate enunciation of a given numeric grouping.

In many cases, however, the performance of these systems are considered deficient, either for not processing adequately natural speech (in the case of speech recognition systems), or for not presenting in its production the expected characteristics of naturally produced speech (in the case of speech synthesis systems). This is due in part to the fact that such systems are mostly based on outdated, non-spontaneous speech data.

Progress has been achieved in that respect as a result of the description of enunciation of natural numbers in various languages, such as German [4], Japanese [5] and French ([6] and [7]). To this date, however, no study has yet described, in a systematic and comprehensive manner, the various acoustic features of the organization of natural numbers in predetermined structures in Brazilian Portuguese.

1.1. Aims of the study

The aim of this study is two-fold: firstly, to segment the number sequence into prosodic sub-groupings in order to investigate the grouping and wording strategy applied to telephone numbers of different length in Brazilian Portuguese. Secondly, to conduct a prosodic analysis in order to determine the typical intonation contours of telephone numbers as spoken by native speakers. The results of such study may serve as valuable information in improving automatic speech recognition and synthesis systems in application connected to telephone numbers in Brazilian Portuguese.

2. Experiment

2.1. Speech materials

The corpus of the present study consists of read-aloud samples of a total of 30 telephone numbers, listed in Table 1. The numbers were all selected from a local telephone book. The numbers were chosen randomly to cover (a) the landline and mobile telephone numbers with eight digits, (b) the special service numbers with three digits and (c) the toll free number with eleven digits.

In order to test a possible relationship between the graphic displays of numbers and the spoken strategy applied in their utterance, the landline numbers were presented in three different types of graphic displays: (i) into two groups of four digits (NNNN-NNNN), (ii) into a group of three digits followed by two groups of two digits each (NNN-NN-NN) and (iii) a group of eight digits (NNNNNNNN). Telephone books contain numbers with these three types of grouping, although the type (i) is the most common.

The 30 telephone numbers were presented in a slideshow at regular interval of seven seconds. The numbers were recorded from 85 adult, native speakers of Brazilian Portuguese (48 women and 37 men). Speech materials were recorded using a minidisc (Sony, MZ-R700) through a digital microphone (Sony, ECM-MS907) placed at 15 centimetres of participants' mouth. Before the recording, the participants did a quick rehearsal with six numbers. They were instructed to utter spontaneously the numbers that were displayed in a computer screen, and were informed that there is no correct way of speaking telephone numbers and that they were not being evaluated in that task.

Table 1: *List of telephone numbers use in the analyses of this study.*

Telephone numbers			
3 digits	8 digits		11 digits
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	

2.2. Segmental analysis

The spoken telephone numbers were extracted from the recordings and archived in *.wav* format using *Praat* as

previously shown [3]. Each spoken telephone number was segmented into intonation units and transcribed orthographically into grouping represented by the digit 1 (unary), 2 (binary), 3 (ternary), or 4 (quaternary), and into wording represented by the letter U (units), D (tens), C (hundreds) or M (thousands) as exemplified in Figure 1.

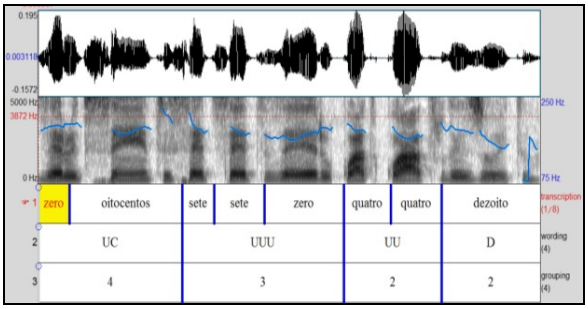

Figure 1: *Praat's textGrid showing the waveform, spectrogram, pitch contour, segmentation and transcription of 08007704418 as spoken by participant woman_2.*

2.3. Intonation contours analysis

Intonation contours of the spoken telephone numbers were extracted through the following method: Firstly, melodic variations of each spoken telephone number were analysed in a semi automatic way with the *Momel/Intsint* script as previously shown [1] in *Praat* (version 5.3.53), using the script' default values (This plugin allows a *Praat* user to convert the pitch contour into a stylized curve, which can be afterwards modified manually according to an auditive evaluation, and finally annotated following the *INTSINT* notation. *INTSINT* is "a transcription system by means of which pitch patterns can be coded using a limited set of abstract tonal symbols, {T,M,B,H,S,L,U,D} (standing for: Top, Mid; Bottom, Higher, Same, Lower, Upstepped, Downstepped respectively)" [8]). Figure 2 shows an example of a telephone number prosodically annotated with the *Momel/Intsint* script.

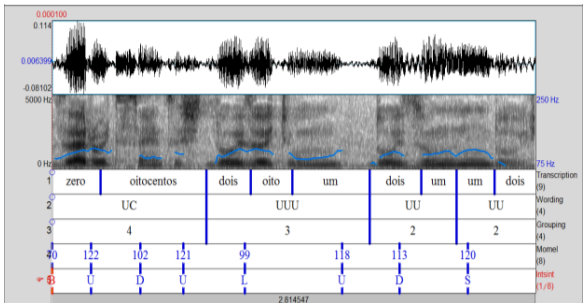

Figure 2: *Praat's textGrid merged showing the waveform, spectrogram, pitch contour, fundamental frequency values, contour variations' coding, transcription and segmentation of 08007704418 as spoken by participant man_69.*

After that, each wording corresponding to a grouping in the spoken telephone numbers was extracted and analyzed with the *Prosodypro* script for *Praat*. In equidistant time intervals, the script was set up to select ten values of fundamental frequency (F0) in the wordings, thereby making it possible to observe adequately intonation contours of wordings of different length through graphic representations. An example is shown in Figure 3.

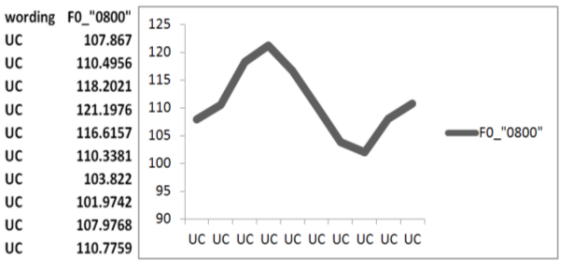

Figure 3: *Graphic representation of intonation contour of "0800" in "08002812112" as spoken by participant man_69.*

3. Results

3.1. Grouping strategy

For eight-digit telephone numbers, the most common grouping strategy is of type "2-2-2-2" in 85% of all cases, which is, four binary groupings.

The grouping strategy with three-digit telephone numbers, in its totality, is of type "3" in 100% of all cases, which is, a ternary grouping.

And with eleven-digit telephone numbers, the grouping strategy is mostly of type "4-3-2-2" in 98% of all cases, which is, a quaternary grouping followed by a ternary and two binary groupings.

The predetermined grouping of the telephone numbers did not seem to influence the grouping strategy adopted by the participants. The grouping strategy of type "2-2-2-2" was preferred in 85% of predetermined grouping of type "NNNN-NNNN", in 80% of predetermined grouping of type "NNNNNNNN" and in 81% of predetermined grouping of type "NNNN-NN-NN".

3.2. Wording strategy

Grouping strategy of type "2-2-2-2" with eight-digit telephone numbers, mostly presented wording of type "$UU_{(1)}$-$UU_{(2)}$-$UU_{(3)}$-$UU_{(4)}$" in 48% of all cases, which is, four binary groupings rendered as units. Other relatively significant wording types are "UU-UU-D-D" in 9%, "D-D-D-D" in 9% and "UU-UU-D-UU" in 9% of all cases.

Grouping strategy of type "4-3-2-2" with eleven-digit numbers, preferentially presented wording of type "UC-UUU-$UU_{(5)}$-$UU_{(6)}$" in 19% of all cases, that is, a quaternary grouping rendered as units and hundreds, followed by a ternary and two binary groupings rendered as units. Other relatively significant wording types are "UC-C-D-D" in 17%, "UC-UUU-D-D" in 16% and "UC-C-D-UU" in 13% of all cases.

Grouping strategy of type "3" with three-digit numbers, mostly presented wording type "C" in 66% of all cases, that is, a ternary grouping rendered as hundreds. Other relatively significant wording type is "UUU" in 34% of all cases.

3.3. Coding of intonation

In spoken telephone numbers with three digits of wording type "C", the most recurrent intonation coding generated by *Intsint/Momel* script in *Praat* is of type "MUD", that is: a Mid

tone, followed by an Upstepped tone and ending with a Downstepped tone.

In telephone numbers with eight digits of wording type "$UU_{(1)}$-$UU_{(2)}$-$UU_{(3)}$-$UU_{(4)}$", the most recurrent intonation coding generated in the first grouping "$UU_{(1)}$" is of type "MUD", that is: a Mid tone, followed by an Upstepped tone and ending with a Downstepped tone. In the second grouping "$UU_{(2)}$", the most generated intonation coding is of type "DU", that is: a Downstepped tone ending with an Upstepped tone. The third grouping "$UU_{(3)}$", mostly, presented the intonation coding of type "UD", that is: an Upstepped tone ending with a Downstepped tone. In the last grouping "$UU_{(4)}$", the most generated intonation coding is of type "UB", that is: an Upstepped tone ending with a Bottom tone..

Lastly, in telephone numbers with eleven digits of wording type "UC-UUU-$UU_{(5)}$-$UU_{(6)}$", the most generated intonation coding in the first grouping "UC" is of type "MUDU", that is: a Mid tone, followed by an Upstepped tone, a Downstepped tone and ending with an Upstepped tone. In the second grouping "UUU" the most generated coding is of type "MD", that is: a Mid tone ending with a Downstepped tone. In the third grouping "$UU_{(5)}$" the most generated coding is of type "UD", that is: an Upstepped tone ending with an Upstepped tone. And in the last grouping "$UU_{(6)}$", the most generated intonation coding is of type "UB", that is: an Upstepped tone ending with a Bottom tone.

3.4. F0 contour

We analyzed variation of the ten selected F0 values at equidistant time intervals in each grouping of spoken telephone number. F0 contours and their patterns were represented.

With eight-digit telephone numbers, Figure 4, 5, 6 e 7 show, respectively, intonation contours of exemplified groupings of type $UU_{(1)}$, $UU_{(2)}$, $UU_{(3)}$ and type $UU_{(4)}$ and their global patterns as spoken by participants.

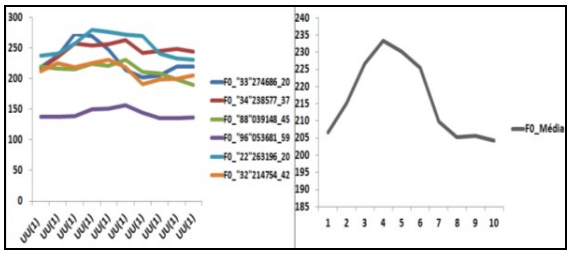

Figure 4: *Intonation contours of exemplified groupings of type $UU_{(1)}$ and their global patterns.*

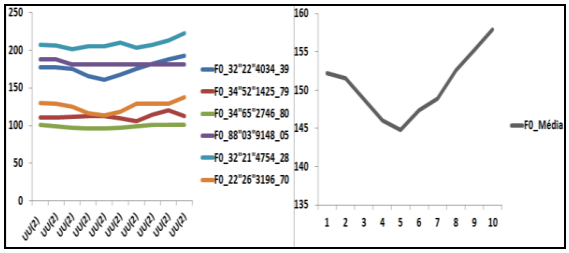

Figure 5: *Intonation contours of exemplified groupings of type $UU_{(2)}$ and their global patterns.*

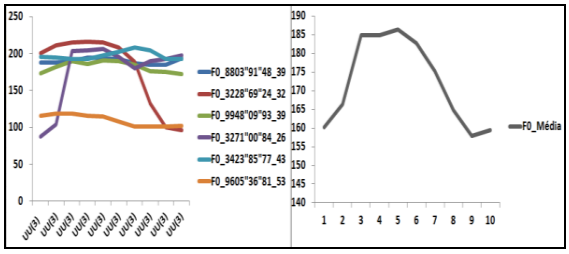

Figure 6: *Intonation contours of exemplified groupings of type $UU_{(3)}$ and their global patterns.*

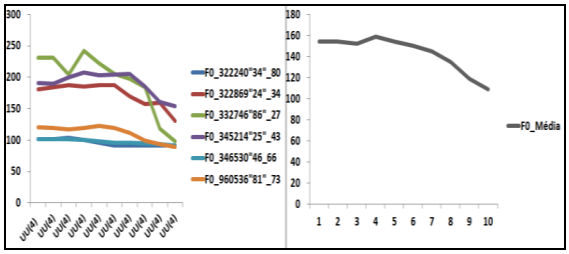

Figure 7: *Intonation contours of exemplified groupings of type $UU_{(4)}$ and their global patterns.*

With eleven-digit telephone numbers, Figure 8, 9, 10 and 11 show, respectively, intonation contours of exemplified groupings of type UC, UUU, $UU_{(5)}$ and type $UU_{(6)}$ and their global patterns as spoken by participants.

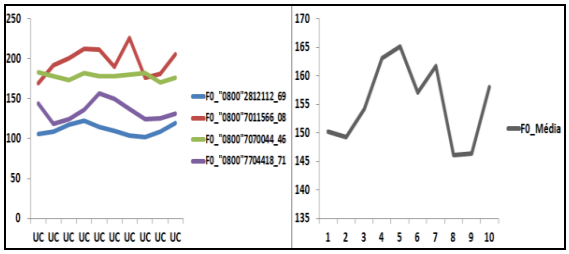

Figure 8: *Intonation contours of exemplified groupings of type UC and their global patterns.*

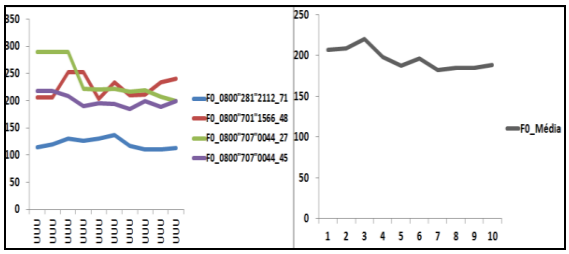

Figure 9: *Intonation contours of exemplified groupings of type UUU and their global patterns.*

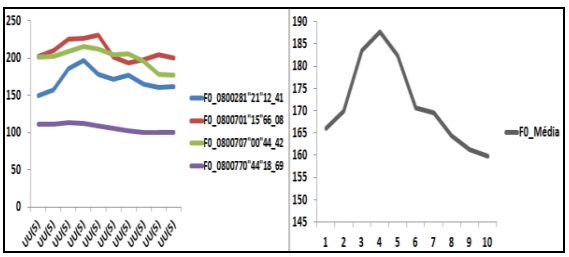

Figure 10: *Intonation contours of exemplified groupings of type $UU_{(5)}$ and their global patterns.*

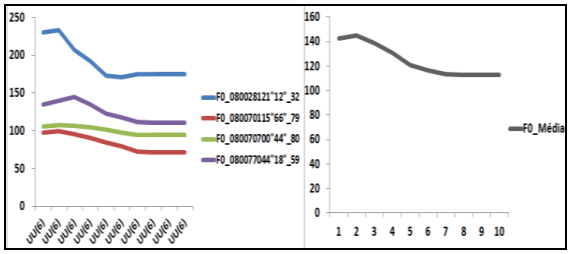

Figure 11: *Intonation contours of exemplified groupings of type $UU_{(6)}$ and their global patterns.*

Figure 12 shows intonation contours of exemplified groupings of type "C" and their global pattern with three-digit telephone numbers as spoken by the participants.

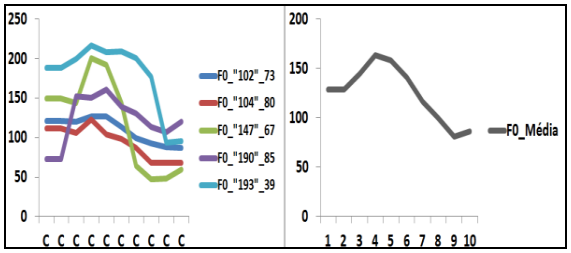

Figure 12: *Intonation contours of exemplified groupings of type "C" and their global pattern.*

3.5. Intonation contour

Results of analyses have shown grouping, wording strategies and intonation contour patterns in spoken telephone numbers in Brazilian Portuguese.

In three-digit telephone numbers, a ternary grouping rendered as hundreds is preferred. As example, the telephone number "190" is spoken as "*cento e noventa*" with the intonation contour shown in Figure 13.

Figure 13: *Intonation contour and corresponding intonation coding of three-digit telephone numbers as spoken by native speakers of Brazilian Portuguese.*

In eight-digit telephone numbers, a four binary groupings strategy rendered as units is preferred. As example, the telephone number "3445 2348" is spoken as "*três quatro*"-"*quatro cinco*"-"*dois três*"-"*quatro oito*" with the intonation contour shown in Figure 14.

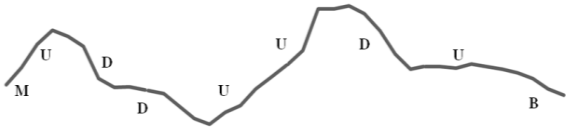

Figure 14: *Intonation contour and corresponding intonation coding of eight-digit telephone numbers as spoken by native speakers of Brazilian Portuguese.*

In eleven-digit telephone numbers, the preferred strategy is a quaternary grouping rendered as units and hundreds, followed by a ternary and two binary groupings rendered as units. As example, the telephone number "08002812112" is spoken as "*zero oitocentos*"- "*dois oito um*"- "*dois um*"- "*um dois*" with the intonation contour shown in Figure 15.

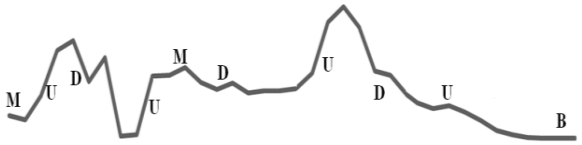

Figure 15: *Intonation contour and corresponding intonation coding of eleven-digit telephone numbers as spoken by native speakers of Brazilian Portuguese.*

4. Conclusion

The main purpose of the present study was to identify intonational patterns of spoken telephone numbers in Brazilian Portuguese. It was demonstrated that intonation alone is enough to characterize prosodically three different types of telephone numbers in Brazilian Portuguese.

A comparison of this study's results with those of [9] shows that the final fall trend reproduce by *Intsint* in Brazilian declarative sentences is also evident in telephone numbers and transcribed by *Intsint*, in this study, with the symbols D (Downstepped tone) and B (Bottom tone) as presented in Figure 13, 14 and 15 respectively. We assume that the intonational patterns of Brazilian Portuguese sentences are used to express the telephone numbers, but the phrasing is adjusted to deliver the sense of grouping.

Other finding of this study is that a similarity in melodic variation exists between telephone numbers in Brazilian Portuguese and those of other languages. The final fall tone of telephone numbers in Brazilian Portuguese is also observed in studies realized on telephone numbers in German [4], Japanese [5] and French [7].

It would of course be interesting to investigate other prosodic features in the characterization of telephone numbers in Brazilian Portuguese, such as intensity and duration. A future investigation involving these features is planned for future analyses. It is also planned to run perceptual tests, based on the intonational patterns identified here in order to find out whether the patterns, applied to synthetic speech, are considered acceptable by Brazilian Portuguese speakers. This research may be regarded as a contribution to the study of spoken numbers in general and to the improvement of automated dialogue systems connected to numbers in particular.

5. Acknowledgements

The completion of this study was made possible with the assistance of *FonUFAL* (Fonetics and Fonology study group of the *Universidade Federal de Alagoas, Brasil*) and the financial support of *CNPq* (*Conselho Nacional de Desenvolvimento Científico e Tecnológico*).

6. References

[1] Hirst, D.J., "A Praat plugin for Momel and Intsint with improved algorithms for modelling and coding intonation", In Proceedings International Conference on Phonetic Sciences, Saarbrücken, 2007

[2] Xu, Y. (2005-2013), "ProsodyPro.praat": http://www.phon.ucl.ac.uk/home/yi/ProsodyPro/, accessed on 19 October 2013.

[3] Boersma, P. and Weenink, D. "Praat: doing phonetics by computer. Versão 5.3.53": http://www.praat.org/, accessed on 20 September 2013.

[4] Baumann, S. and 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.

[5] Amino, K and Osanai, T., "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 17-21, 2011.

[6] Bartkova, K. and Jouvet, D., "Selective prosodic post-processing for improving recognition of French telephone numbers", In Proceedings of The 6th Eurospeech. Budapest, Hungary, 1999.

[7] Martin, P. La prosodie. 2007. Available in: < http://www.linguistes.com/phonetique/prosodie.html>. Accessed on 23 February 2014

[8] Hirst, D.J., Di Cristo, A. and Espesser, R. (2000). Levels of representation and levels of analysis for the description of intonation systems. Prosody: Theory and Experiment. Kluwer Academic Press.

[9] Reis, c. & Celeste, l.c. Análise entonativa formal:INTSINT aplicado ao Português. Journal of of Speech Sciences, Vol 2, Iss 2, Pp 3-21 (2012). Available in :< http://www.journalofspeechsciences.org/index.php/journalofspeechsciences/article/view/49/47>. Accessed on 23 Febrary 2014.