Pitch level variation in French: categorical or gradient?

David Le Gac[1], *Hiyon Yoo*[2], *Katarina Bartkova*[3]

[1]Rouen University, Laboratoire DySoLa, France
[2]Paris-Diderot University, Laboratoire de Linguistique Formelle, Sorbonne-Cité, France
[3]Nancy University, ATILF, France

david.legac@univ-rouen.fr, yoo@linguist.jussieu.fr, katarina.bartkova@atilf.fr

Abstract

This paper reports the results of an experiment on the question of whether the realizations of f0 variations at the end of the final IP are categorical or gradient. We conducted an imitation task with resynthesized stimuli where the final pitch height was varied in steps of one semi-tone. Results are ambivalent, since both strategies are possible. However, we argue that there is enough evidence for establishing the existence of at least three pitch categories in French.

Index Terms: intonation, imitation task, French, categorical, continuous, pitch level, pitch range

1. Introduction

A central question when modeling intonation is how variations of fundamental frequency (f0) are divided into phonological distinctive units. For instance, what is generally accepted within the Metrical and Autosegmental (MA) framework ([10], [13]) is that there are two tonal categories L(ow) and H(igh). However, these two categories are not uncontroversial since other authors have proposed systems with more than two tonal levels (cf. [10] for a review).

For French, the question of the number of distinctive pitch levels remains an issue, in particular at the final-IP position. For instance, [4] and [8], who follow the standard MA framework describe French intonation with the two H and L tones; [15]'s MA tone grammar, however, generates three contrasting pitch levels, L%, H% and Ø%. [3] and [18] oppose to the infra-low level two higher levels for the final IP position, the high and supra-high levels supporting the existence of at least three levels in this position. [3] proposes also the feature "+" that could be interpreted as an additional level. Finally [12] propose a model with four different tonal levels for the final IP position (namely B–, B, H, H+).

One way of addressing the question of the number of distinctive pitch levels is to tackle this issue experimentally and investigate whether pitch contrasts are gradient or categorical in perception and/or production.

In the MA framework, height variations of the L and H tones are considered as paralinguistic gradient variations of pitch range. This gradient treatment of pitch range is supported by experiments performed by [11] where subjects could successfully produce ten overall pitch ranges without grouping them into smaller categories.

Several studies have applied the categorical perception (CP) method to different intonational contrasts. [9] examine the contrast between "normal" and "emphatic" accent peaks in English. They come up with ambivalent results: they find a well formed S-shaped identification curve but the discrimination task function does not exhibit a clear peak at the inferred category boundary. They conclude that "normal" and "emphatic" accent peaks in English can be considered as categorical in production, but not in perception. [17] and [19] test the categorical status of high and low boundary tones in Dutch and German respectively. The overall results are neither consistent with continuous perception nor with classical CP. In particular, the discrimination curves show either a plateau or two peaks. This leads the authors to suppose that there might be a third ('hidden') category between the falling statement and the sharply rising question. Applying a similar approach for Catalan, [20] presents results supporting the idea that the pitch height represents the phonetic basis of a phonological distinction between yes-no questions and wh-questions.

[14] use an imitation task in order to study whether tonal alignment of rise-fall-rise in English lies on a categorical or a gradient distinction. Their results support a binary distinction for tonal alignment: there is a systematic deviation of the peak alignment between the stimuli and the responses of the subjects, and the distribution of the peak delays in the imitations is bimodal. In two other studies running imitation tasks, [5] and [16] come up with clear results showing that there is a categorical effect of graded stimuli on pitch movements (pitch range variation and f0 shifts). However, in another study where an imitation task was conducted in order to test pitch range, [6] obtains contradictory results showing that speakers were able to produce continuous rather than categorical responses. However, [6] argues that it is not because speakers were able to produce a continuum that categories do not exist, and suggests that stimuli express graded variations within a single phonological category, and these variations can be reproduced by the speakers.

For French, the only study dealing with the categorical vs. continuous distinctions for intonation can be found in [15]. [15] applies the paradigm of categorical perception in order to test pitch contours at the end of the final IP in French. The results are ambivalent since the identification task curve for the contrast between the pitch levels H*Ø and H*H% has the shape of a categorical distinction, but the discrimination peak is not properly aligned. The overall results do not provide strong evidence for distinguishing three distinct categories.

In this paper, we investigate the pitch height variations at the end of the final IP in French by conducting an imitation task. Our aim is double: First, we test if the realizations of the f0 variations are gradient or categorical; second, if the realization is categorical, we aim to provide the number of levels that are distinctive.

2. Experimental Procedures

2.1. The choice of the experimental approach

[15]'s ambivalent results as well as those of [9], [17] and [19] lead those authors to suggest that the 'classical' CP paradigm seems not to suit the issue of categorical vs. gradient distinction in intonation. This question is also discussed in [2], [7] and [14]. As put forward by [14], identification tasks automatically force responses into categories; such tasks cannot reveal whether categories exist but only where the

boundaries between categories are. These tasks are thus to be used when the phonological categories are uncontroversial and are related to the issue of the status of the linguistic description in the psychological system. For the present study, we chose to conduct an imitation task in line with [6], [5] [16] and [14], which seem to be the best method for testing putative phonological oppositions for intonational categories.

2.2. Stimuli

First, we recorded natural productions of a male speaker who pronounced the sentence “Elle est là” ‘she is here’. Different contexts were given in order to obtain final rise and fall contours with the different possible pitch levels. We chose as the initial stimulus the sentence with the prefinal and final pitch values that corresponded the most to the average prefinal and final f0 values for all pronounced utterances. This natural sentence was then resynthesized by varying the final pitch height in steps of one semi-tone (the resynthesis was done using the PSOLA algorithm with the Praat software package ([1]). A flat f0 was preserved for the first two syllables. We obtained 26 stimuli that were randomized in blocks.

2.3. Participants

Subjects were seven native speakers of French, staff and students at a French university (3 three males and four females) with no hearing problems. One of them MRG is a PhD student in phonetics and has an experience as an actor. The six other participants are not accustomed to phonetic experiments. None of the participants was informed of the purpose of the experiment.

2.4. Procedure

Each subject performed the imitation task in two sessions, for a total of 14 different blocks. Each session was preceded by a training session containing only nine stimuli. We gathered a total of 2548 tokens for analysis (364 tokens per subject). The recording software used in our study was developed by the LORIA research center (Nancy, France). The original version of the software was adapted to the needs of the experiments carried out so that subjects taking part in the experiments managed time as well as the recording software: there were no time restrictions and they could listen to the stimulus as many time as they wanted, listen to their own production and rerecord their response as many times as they wanted.

Since it was an imitation task, subjects were asked to repeat as exactly as possible the stimulus they heard. They were informed that the sentence they would hear would be the same for the whole experiment, varying only in intonation. The experimenter could intervene only during the training session and between two blocks. Each block lasted 4min 50s in average (for a range from 3 to 10 minutes).

We measured the minimum or the maximum of the f0 value on the final vowel depending on the direction of the slope (rising or falling), since these values are considered in the current theories of French intonation as the relevant points distinguishing the different pitch categories. These values were then converted into semi-tones (ST), in order to allow comparison between male and female speakers, and to normalize the variation between low and high values within a same speaker. We used the mean of all values in hertz of a speaker as the referent value for the ST calculation. Using a ST scale is also in adequation with the stimuli, which were obtained by varying the final pitch height in steps of one ST.

3. Results

In this section, we present the scatterplot of median f0 height and the histogram for each speaker. The median f0 values for responses in semi-tones “ST” are plotted on the vertical axis against the stimuli on the horizontal axis, with first and third quartiles indicated. The histogram expresses the frequency distribution in absolute number of f0 values in ST for each stimulus (i.e. 26 bars).

Results show that subjects do not present a unique behavior; three groups of speakers have to be distinguished.

1. The realizations of the speaker MRG are remarkably close to the initial stimuli, as can be seen in Figure 1. The histogram shows a quasi-even distribution. From MRG’s data, we can say that reproducing an exact continuum is possible. However, what distinguishes MRG from the other speakers is that he is a PhD student in phonetics and that he has an experience in acting. Finally, compared to the other speakers, he took twice the time to accomplish the task than the other subjects. The rather exceptional results of MRG can be explained by taking these facts into account.

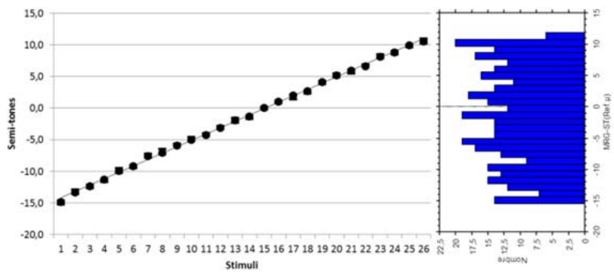

Figure 1: *speaker MRG; plot of median f0 height (semi-tones) with first and third quartiles for responses to each stimulus. Histogram of f0 height.*

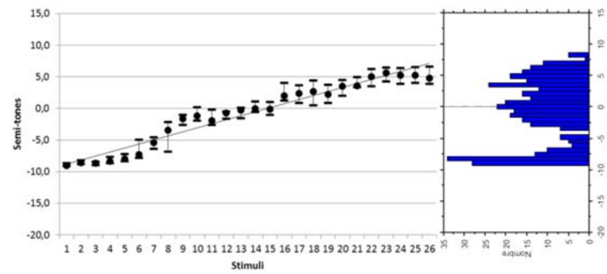

Figure 2: *speaker FZJ (plot of median f0 height and histogram)*

2. On the opposite side, FZJ’s data support the existence of at least three distinct categories. The histogram of Figure 2 presents a trimodal distribution, with three distinct peaks around –8, 0 and 3 STs. Examination of the median plot shows that there are important deviations of the responses compared to the initial stimuli, and suggests that even four different clusters can be formed: the first one on the lower arm of the curve grouping responses of stimuli 1 to 5, a second one in the middle between 9 and 15; responses 6-8 are transitory responses (with large variations in quartiles). Finally, two further clusters may be defined in the higher arm of the curve, between 16-19 and 22-26 with two ambiguous and transitory responses (20-21). However, the distinction between these two categories is not clear enough because the third quartiles of the third cluster overlap with the first quartiles and even with median values of the fourth cluster.

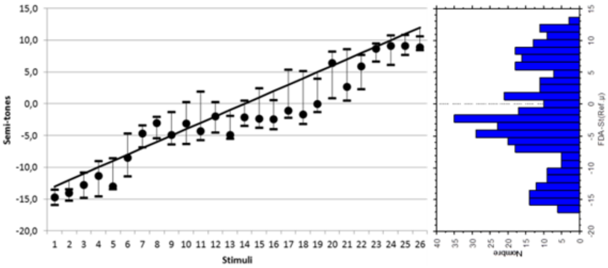

Figure 3: *speaker FDA; plot of median f0 height (semi-tones) with first and third quartiles for responses to each stimulus. Histogram of f0 height.*

Data for FDA also support the idea of 3 distinctive categories (Figure 3). The median plot shows 3 clusters: (i) responses to stimuli 1-5; (ii) 7-16 and (iii) 23-26. The histogram shows also a trimodal distribution (around −15, −3 and 8ST), the prominent mode being around −3 ST, which corresponds to the large plateau of the second cluster in the plot of median values. Response to stimulus 6 is transitory, with broader quartiles. For both FZJ and FDA, the zone that covers stimuli 6 to 8 can be considered a transitory zone marking a boundary for a "low" category.

For responses to stimuli 17 to 22, there is a broad "zone" with variation and instability of the median value in speaker FDA's data. The results suggest that these stimuli were treated as ambiguous between the second and the third category.

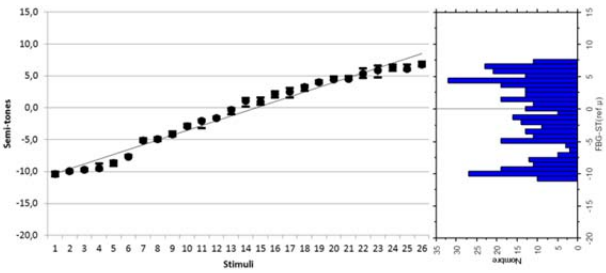

Figure 4: speaker *FBG (median f0 plot & histogram)*

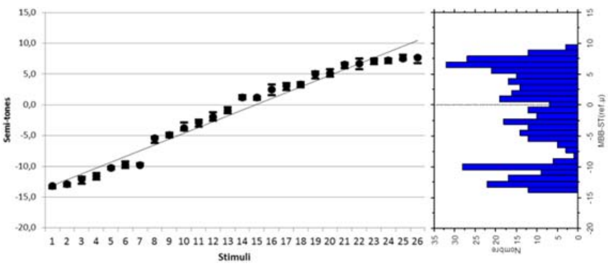

Figure 5: *speaker MBB (median f0 plot & histogram)*

3. Results for the other speakers (FBG, FLM, MBB and MSJ, Figures 4 to 7) are more controversial. When looking at their scatterplots, they are closer with MRG's median values with no cluster formation as levels like in FZJ and FDA's data. Data suggest that speakers MBB, MSJ, FLM, and FBG succeed in closely imitating the stimuli and generally with little variation. However, the patterns of their corresponding histograms differ clearly from MRG's histogram and reveal distinct modes - as those of FZJ and FDA - suggesting the existence of distinctive pitch height.

The existence of pitch height can be further supported by visible "jumps" disturbing the linearity of scatter plots. We define "jumps" as a difference between two immediate median values of at least 2 STs and with no overlap between the third quartile of the first value and the first quartile of the second value. Moreover, the location of these jumps corresponds to a valley in the histograms. Thus, we propose to interpret these jumps associated to valleys in histograms as boundaries of pitch categories.

Thus, for all speakers, we observe a first mode around −10ST, forming a first cluster which corresponds to a low pitch level - as was observed for FZJ and FDA. Furthermore, for speakers FBG and FBB, there is a substantial jump between stimuli 6 and 7 (cf. valleys in histograms around −6ST for FGB and −8ST for FBB). This first cluster encounters the tone category Low, generally associated with assertion. As for the higher values, the number of distinctive modes differs from a speaker to another.

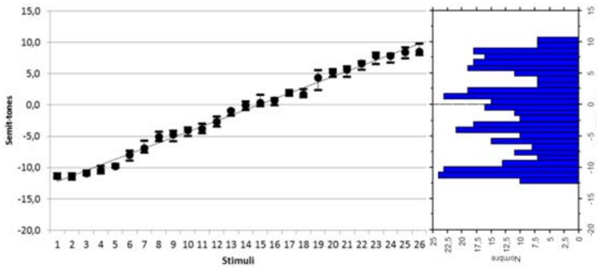

Figure 6: speaker *FLM (median f0 plot & histogram)*

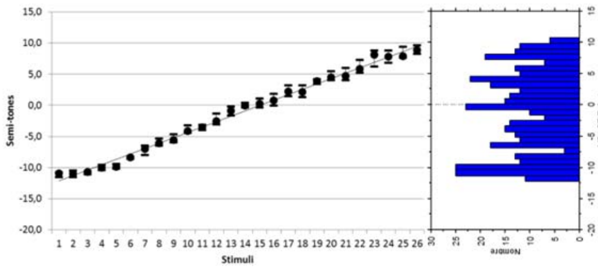

Figure 7: *speaker MSJ (median f0 plot & histogram)*

For FBG (Figure 4), the histogram suggests a bimodal distribution, with peaks around −10ST and 4ST. The histogram for MBB (Figure 5) reveals two peaks around −10ST and 7ST. The association of jumps (between responses to stimuli 7/8 and 13/14) with valleys (−8ST and 0ST) makes us propose a third category where the boundaries are located around responses stimuli 7/8 and 13/14, the peak being around −3ST. For FLM (Figure 6), the histogram shows four modes, with peaks around −11ST, −4ST, 2 and 7ST. The scatterplot does not present clear clusters as in FZJ and FDA but global progression is less linear than MRG. In fact, responses to stimuli 6 and 7, 12 and 13 and 19 seem to be transitory values defining four different categories. For MSJ (Figure 7), the histogram and the scatterplot are difficult to interpret, with the exception of the low values. They suggest several preferred values around −10, −6, −1, 4 and 8 ST, but without clear distinctions between pitch categories.

4. Discussion

The imitation task we conducted does not provide homogeneous results. For speakers FZJ and FDA, median f0 plots reveal a pattern in steps, to which we can associate a trimodal distribution in histograms. These results can only be explained if we postulate at least three categories underlying pitch height. The pattern of the median plots is close to the results obtained by [14] allowing the authors to conclude tonal alignment is categorical in English, by using the same experimental paradigm. These results are also consistent with [17], [19] and [15]'s studies, which suggest a third category between the L% and H% tones.

On the other hand, five speakers out of seven could track the values for the stimuli with little variation. These results may find three explanations.

1. Our data suggest that pitch height variations are gradient rather than categorical, at least for some of the speakers. They comfort the general idea that there can be a gradient modulation of pitch range in line with results found by [11] and [6] and discussed by [7]. It would mean that there are no distinctive tonal categories, even between a Low tone and a High tone since the stimuli of our experiment are built from a spectrum that goes from low to extra-high values. However, it is unlikely that at least these low pitch values that generally correspond to the low tones (or low "register") in traditional theories do not form a category. In addition, we proposed that jumps in median values combined with valleys and peaks in histograms constituted cues for tonal categories. In particular, for speakers FBG, MBB, and FLM, a clear cut can be found systematically between a low category and higher categories, and also within higher categories.

2. It can be proposed, therefore, that tonal categories do exist for these speakers, but that each category may constitute a gradient dimension with a preferred value (cf. [7]).

3. A third and more likely interpretation is that gradient imitation is due to the task itself: speakers who produced the continuum of stimuli have competence in imitating fine phonetic details of another speaker (an idea put forward by [14]). In other terms, being able to reproduce a continuum does not mean that categories do not exist (see also [16]). Moreover, our data show that performance is related to time devoted to the experiment; thus speaker MRG who realized the best performance spent twice the time than the others per block of 26 stimuli (8 min 17, the average time for all speakers being 4 min 50).

FZJ and FDA's data corroborate indirectly the idea that the gradient realization is due to the task. What our results show is that FZJ and FDA are not as performing imitators as the other speakers; thus, their productions were spontaneously brought to their system of tonal categories. Moreover, the fact that their realizations vary more than for the others supports this idea that imitating the pitch variations of another voice is a difficult task and that their categories prevail over the target values to perform.

5. Conclusion

Although speakers were able to reproduce pitch height variations with great precision, we argued that significant deviations between the stimuli and the responses of the subjects can only been explained if we admit the existence at least three pitch categories in French. We suggest that gradient productions by speakers are due to the imitation task; in particular, further experiments with more restricted instructions should be carried out.

If we admit the existence of at least three categories, then an intonation model that comprises only two tones such as [4] and [8] for French appears not sufficient to account for our data. Models based on a three- or four-way opposition such those of [15], [12] or [18] would be more adequate to account for our data.

6. Acknowledgments

The authors wish to thank the participants, E. Wander for his corrections and remarks and the group Parole of the LORIA laboratory, and particularly V. Colotte for kindly allowing us to use and adapt the recording software (CorpusRecorder used in Intonal project) and D. Jouvet, for the automatic segmentation of the speech data, and the computation of the f0 parameters.

7. References

[1] Boersma P. & Weenink, D. "Praat, a system for doing phonetics by computer". Glot International 5(9/10): 341-345, 2001

[2] Cummins, F., Doherty, C., & Dilley, L. "Phrase-final pitch discrimination in English". In Proceedings of Speech Prosody 2006, Dresden, Germany, 2006,

[3] Delattre, P. "Les dix intonations de base du français", The French Review 40, 1-14, 1966

[4] Di Cristo, A. "Intonation in French". in Di Cristo, A. & Hirst, D. (eds), Intonation Systems: a Survey of Twenty Languages, Cambridge University Press, 195-218, 1999

[5] Dilley, L. & Brown, M. "Effects of pitch range variation on F0 extrema in an imitation task". J.of Phonetics, 35, 523-551, 2007

[6] Dilley, L., "Pitch range variation in English tonal contrasts: Continuous or categorical?" In Proceedings of the 16th ICPhS, Saarbrücken, Germany, 1153-1157, 2006

[7] Gussenhoven, C. "Experimental approaches to establishing discreteness of intonational contrasts", in S. Sudhoff & al. (eds), Methods in Empirical Prosody Research, MdG, 321-334, 2006

[8] Jun, S.-A. & Fougeron, C. "A Phonological model of French intonation". In A. Botinis (ed.) Intonation: Analysis, Modeling and Technology. Dordrecht: Kluwer, 209-242, 2000

[9] Ladd, D.R. & Morton, R. "The perception of intonational emphasis: Continuous or categorical?", J. of Phonetics 25, 313-342, 1997

[10] Ladd D. R., Intonational Phonology, second edition. CUP, 2008

[11] Liberman, M. & Pierrehumbert, J. "Intonational invariance under changes in pitch range and length". In: Aronoffand & Oehrle (eds.): Language sound structure: Studies in phonology presented to Morris Halle. MITPress, 157–233, 1984

[12] Mertens, P., "L'intonation" in Blanche-Benveniste, C. et al (eds), Le français parlé, Paris: Éditions du CNRS, 159-176, 1990

[13] Pierrehumbert, J. & Beckman, M. Japanese Tone Structure, Linguistic Inquiry Monograph 15, MIT Press, Cambridge. 1988

[14] Pierrehumbert, J. & Steele, S.A. "Categories of tonal alignment in English", Phonetica 46, 181-196, 1989

[15] Post, B. Tonal and phrasal structures in French intonation The Hague: Holland Academic Graphics, 2000

[16] Redi, L. "Categorical effects in the production of pitch contours in English", in Proceedings of the 15th ICPhS, Barcelona, Spain, 2921-2924, 2003

[17] Remijsen, B. & Heuven, V. J. van. "Gradient and categorical pitch dimensions in Dutch: diagnostic test". In Proceedings of the 14th ICPhS, San Francisco. 1865–1868, 1999

[18] Rossi, M. L'intonation, le système du français : description et modélisation. Ophrys, 1999

[19] Schneider, K. & Lintfert, B. "Categorical perception of boundary tones in German". In Proceedings of the 15th ICPhS, Barcelona, Spain, 631–634, 2003

[20] Vanrell Bosch, M. M., "A scaling contrast in Majorcan Catalan interrogatives". In Proceedings of Speech Prosody 2006, Dresden, Germany. 2006