ISCA Archive
http://www.isca-speech.org/archive

Speech Prosody 2008
Campinas, Brazil
May 6-9, 2008

The native language influence on perceptual Intrinsic pitch: Cross-linguistic data from German, Italian, Portuguese, and Spanish

Daniel Pape

ZAS (Centre for general linguistics)
Berlin, Germany
pape@zas.gwz-berlin.de

10.21437/SpeechProsody.2008-165

Abstract

It is generally acknowledged that Intrinsic Pitch differences are found for Germanic languages. In a previous study we could give evidence for a strong cross-linguistic difference when examining Intrinsic pitch for the Romance language Italian. To examine the possibility of a stable cross-linguistic difference, a perception experiment was designed to examine the existence of Intrinsic pitch in Romance languages in comparison to Germanic languages. Since Intrinsic pitch values were found with respect to differing musical education, the experiment was conducted separately for professional musicians on the one hand and non-musicians on the other hand.

Different response patterns were found when comparing the Germanic and the Romance languages: Nearly 50 % of the Spanish and Portuguese listener showed an insensitivity for the Intrinsic pitch task, i.e. they judged the pitch of the vowels mainly on the voice quality difference but not on the F0 difference. This insensitivity clearly classifies Intrinsic pitch as a language-specific phenomenon.

1. Introduction

On the speech production side, IF0 describes the dependence of the fundamental frequency on vowel height: Other things being equal, high vowels have a higher F0 compared to low vowels. On the speech perception side, Intrinsic pitch describes the dependence of the perceived vowel pitch on the height of the presented vowel: Higher vowels are perceived with a lower pitch impression compared to low vowels.

Concerning the reasons for Intrinsic pitch on the perceptual side, in the literature it is argued that Intrinsic pitch serves either as:

- A compensation mechanism for the IF0 differences in speech production with the aim to avoid disturbance in the listeners' parsing of the prosodic F0 contour [1] or
- A psychoacoustic phenomenon, that causes a perceived pitch difference due to masking and other effects of the human auditory processing system [2].
- Since the distance between F0 and F1 can be actively used by the listener to change the perceptual impression of the corresponding vowel openness [3], other researchers claim that Intrinsic pitch differences are used for this purpose and therefore to enhance existing vowel contrasts.

All these theories are faced with the drawback that the conducted experiments only examined Germanic languages (German or English) but did not extend the results to general prosodic or psychoacoustic phenomena.

However, the examination of perceived pitch differences between high and low vowels for other than the Germanic languages (e.g. Romance languages like Spanish with only a five vowel contrast) is still missing. It is possible that due to the missing overlap of articulatory settings (and thus formants) of vowels, Intrinsic pitch would not play the same role as for Germanic languages with their more crowded vowels systems, since the different vowels in Romance languages can be regarded as already highly perceptually distinct. It has still to be given evidence if Intrinsic pitch as a perceptual phenomenon is at all existent in other languages, since both the perception of the parsing of articulatory gestures and the perception of prosody and microprosody could differ to a large amount when comparing different languages.

Interested in these Intrinsic pitch differences between different languages and language families, the studies [4] and [5] examined vowels differing in vowel roundedness and height comparing German listeners to both Catalan and Italian listeners. In these studies it was found that Intrinsic pitch is not as stable in the Romance languages compared to the Germanic language, and is often completely missing. Furthermore, in contrast to the expected result that musical education of the listeners would improve accuracy of pitch differences (which was shown e.g. in [6] for musical signals and could be given evidence for Intrinsic pitch for German listeners in [5]), the musically educated listeners of the Romance languages did not show the expected result.

Given these discrepancies, the current study aims to extend the previous studies by examining Intrinsic pitch for additional Romance languages. The languages examined in this study include German as a representative of the Germanic languages and Italian, Portuguese and Spanish as Romance languages.

The study will therefore shed light on the hypothesis whether Intrinsic pitch may be language-dependent consistently and thus not as universal as it is regarded up to now.

Since the studies [1] and [2] did not control certain parameters like musical education, for the current study a cross- linguistic valid experiment was designed with the same fixed parameters and factors for all languages to study the existence of Intrinsic pitch in non-Germanic languages compared to Germanic languages.

In the following, we concentrate on giving an idea of the main cross-linguistic differences and the effect of musical education. The actual data and statistics for the remaining listeners (see definition of *pitch-sensitive* versus *insensitive* listeners later on) will be presented elsewhere.

1. Methods

The complete perceptual experiment is conducted in two separate parts: (1) a pre-study to test the listeners general F0 difference threshold and (2) the Intrinsic pitch differences for both German stimuli and stimuli of the listeners' native language. Here, only the Intrinsic pitch differences (2) are presented.

During the experiment, vowel pairs with differing F0 are presented (vowels /i a/ to test for maximal vowel height differences). The first experiment used native German vowel pairs and the second experiment vowels of the listeners' native language. All stimuli were cut from a naturally spoken sentence in nasal context in stressed position. The stimuli were time-normalized to an ambiguous vowel length to eliminate possible differences caused by duration differences and showed a flat F0 contour.

The test design was 2AFC (with one stimuli pair). The listener had to decide whether she thinks that the first or second vowel in the pair would show the higher pitch. The order of the vowels presented (/i/ and /a/) was randomly chosen. The reference F0 was 120Hz, and fundamental frequencies were shifted ±10Hz in 2.5Hz steps.

In the experiment, for each vowel pair the listener is forced with the task to give the attribute "higher in pitch" to one vowel out of a pair differing both in vowel quality (/i/ versus /a/) and F0 (±10Hz). No feedback was provided to the listener, since there was no objective right/wrong decision. Instead, the interesting parameter was the listeners judgment of the his/her *perceived* vowel pitch differences

1.1. The listeners

The test was conducted with German listeners and Italian, Portuguese and Spanish listeners to examine Romance languages listeners. Furthermore, the experiment was separately conducted with two distinct groups of listeners: Professional musicians on the one hand and listeners who did not play an instrument on the other hand. The number of listeners was about 30 for each group (i.e. 30 German non-musicians, 30 German professional musicians, 30 Italian non-musicians etc.).

Concerning the population of non-musicians, students from different phonetics departments (Berlin, Munich, Kiel) were chosen for the German language. The Portuguese listeners came from the natural sciences department of the university of Florianopolis (Brazil) and the Spanish listeners from the natural sciences department of the university of Valdivia (Chile) – thus, for Spanish and Portuguese we test the South American dialects. For all Romance language we made sure that the listeners were neither educated in English nor in German to avoid possible L2 interference when judging the German stimuli.

The professional musicians were recruited from the music conservatories of the corresponding cities. Only musicians having a degree of the music conservatory were tested to guarantee an equal level of auditory and musical education among all languages.

2. Results and Discussion

Please note that in this proceeding paper only the overview of the used statistical analyses is given. The aim is to give an overview how the perceptual data looks like. Of course, this complex experiment was statistically analysed in various manners: The complete (repeated measurements) MANOVAs are presented for example in my published PhD dissertation.

First, probit analysis is applied to fit the curves of individual subjects. The dependent measure is the obtained probability of the between-vowel F0 difference, for which subjects judged the vowel /i/ to be perceived higher in pitch on 50% of the opportunities. The corresponding F0 value defines the intrinsic pitch difference for each subject. A subsequent t-test extracted whether the intrinsic pitch values significantly differ from 0Hz, with 0Hz meaning no between-vowel intrinsic pitch differences.

Strong differences are found when thoroughly examining the response pattern across populations and languages. These differences can be mainly manifested in what is defined here as pitch-*sensitivity* versus pitch-*insensitivity*: Independent of the presented between-vowel F0 differences, the pitch-insensitive listener always judges one of the vowels higher in pitch, and therefore more judges vowel quality than F0 differences. For example, inspecting the /i a/ response pattern, the listener shows a ceiling effect for all presented F0 differences (-10Hz to +10Hz), instead of the expected monotone rising function. This monotone rising function would be expected based on the experimental setup, and is shown in figure 1). Exhibiting the described pitch-insensitivity, the listener e.g. judges the vowel /i/ *always* higher than /a/, in complete independence of the between-vowel difference – i.e. even though /a/ is presented with a F0 increase of +10Hz.

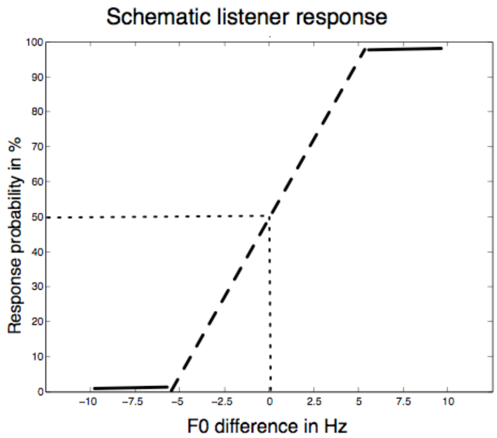

Figure 1: *Schematic view of the listener responses in dependence of the presented F0 difference between the two vowels in the presented pair.*

2.1. Pitch-insensitivity as a cross-linguistic phenomenon

To obtain a better measure of the insensitivity occurrence across all languages, figure 2 shows for each language the percentage of listeners whom are confirmed pitch-insensitivity (according to the above defined criteria, and only when both the German vowel conditions and the native vowel conditions agreed in pitch-insensitivity).

As can be seen, the percentage is lowest for German with 6%. In contrast, for the languages Spanish and Portuguese the highest percentages are found: Nearly half of all listeners are pitch-insensitive in these languages. Italian is found in intermediate position (with 23%). Thus, a clear difference between German and the other languages is found here, with a high percentage of insensitivity for the Romance languages. Further, it seems that the intermediate position for Italian fits

well with the idea that a research stream gives for Italian (see e.g. the results of [7]: The authors define Italian as being a mixture between stress-timed languages (like German) and syllable-timed languages (like Spanish). This intermediate position also probably results in intermediate position for fundamental frequency use with respect to linguistic stress and accent, and could be the reason for the intermediate position in the intrinsic pitch task.
When examining the occurrence of pitch-insensitivity in dependence of musical education, it is found that for all languages only the non-musicians are pitch-insensitive, i.e. that for all languages none of the professional musicians showed pitch-insensitivity for the /i a/ contrast.

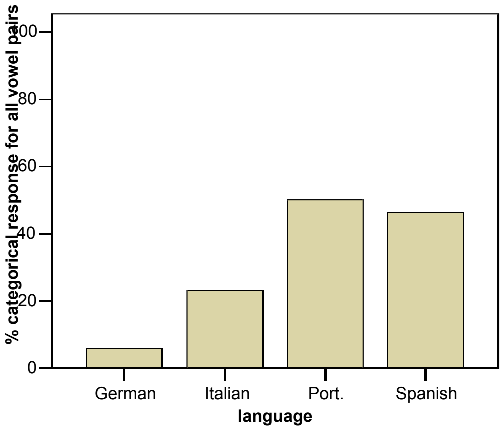

Figure 2: *Percentages of the listeners showing pitch insensitivity response pattern (see text for explanation) – these subjects are all non-musicians.*

2.2. Listener response behavior

Since figure 2 only shows a general trend, it is important to see the specific interaction and dependence of the varied experimental factors. Thus, the aim is to obtain a general idea of the probability function distribution with regard to language, musical education. Figures 3 clarifies these ideas by means of regression lines for pitch-sensitive listeners versus pitch-insensitive listeners for all examined languages. The advantage of using regression lines is based on the idea of visualizing general trends and between-group differences.

Due to the fact that all professional musicians are pitch-sensitive, the regression lines for pitch-sensitive listeners are nearly identical across all languages. They show the same expected behavior, with a rising pattern from 0% at –10Hz towards 100% at +10Hz. With respect to language, the regression lines for the pitch-insensitive listeners show clear differences, with the regression line for German being steepest. No difference in rise is seen when comparing Italian, Portuguese, and Spanish. For all languages, the regression lines for the pitch-insensitive listeners tend to be located more in the lower part of the figure. In addition to the shallow rise, this indicates a tendency that the vowel /a/ is heard lower than /i/, thus manifesting a general intrinsic pitch trend. Although here this trend is mostly independent of the between-vowel F0 difference, it can be noted being absolutely in line with the classical intrinsic pitch results. Interestingly, it can be seen that the steepness of the regression lines is in the order German > Italian > Spanish > Portuguese.

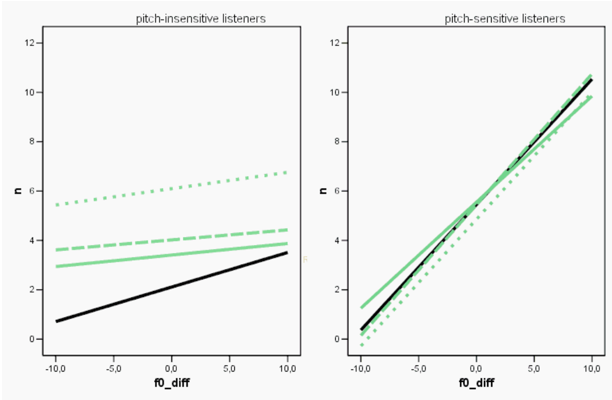

Figure 3: *Pitch-sensitivity for the German vowel condition: Shown are the regression lines for the absolute listener judgments (/i/ heard higher in pitch than /a/, n equals the number of trials with n_{max}=12) versus between-vowel F0. The left panel shows the data for pitch-insensitive listeners, the right one for pitch-sensitive listeners. Black solid lines show results for the German language, whereas all green lines indicate Romance languages (solid Italian, dotted Portuguese, and dashed Spanish).*

3. Conclusion

The previous section shows that the response patterns are clearly different when comparing Intrinsic pitch for German listeners to of Romance language listeners. It is given evidence that Romance listeners of Spanish and Portuguese are highly insensitive to the presented F0 differences: Nearly 50% of all subjects did not show the expected pitch response behavior - thus it was not possible to extract a consistent Intrinsic pitch value in the presented F0 range of ±10Hz. It can be ruled out that the F0 range causes this insensitivity: In the literature it is shown ([1]) that ±8Hz are completely sufficient to generate Intrinsic pitch differences for English and German listeners. Therefore, it is assumed that the insensitivity of the Romance listeners is due to a language-specific phenomenon. Due to the small vowel inventory it could be the case that the perceptual cue vowel pitch is not that important as in other languages. Therefore, it is not necessary to use it as an additional cue to e.g. robustly perceive vowel openness. However, further experiments – mainly identification tests with varying F0 parameterization – will shed light to the question whether in fact Intrinsic pitch is completely absent in Romance languages. Concerning the Germanic languages, it is still to be given evidence to which amount the cue intrinsic pitch is used - or can be used – by listeners to achieve a more robust identification or compensation for production differences.

With respect to the role of intrinsic pitch for prosodic parsing (as presented by Fowler), the here-found results completely rule out the possibility of compensation for Romance languages: In order to use compensation strategies to achieve stable micro-prosodic parsing in the Romance languages, the listeners are forced to robustly identify intrinsic pitch differences. This was proven for English and German listeners, but the here-presented results for the Romance language listeners do not support this idea. As a result, the lack of Intrinsic pitch aims at a substantial difference in prosodic parsing and calls for further examination of the role of F0 for speech production, speech perception and the link between the two areas of research.

4. References

[1] Fowler C.A. and Brown J.M. 1997. Intrinsic f0 differences in spoken and sung vowels and their perception by listeners, *Perception and Psychophysics*, vol.59, no.5, pp.729-738.

[2] Stoll G. 1984. Pitch of vowels: Experimental and theoretical investigation of its dependence on vowel quality, *Speech Communication*, vol.3, pp. 137-150.

[3] Traunmüller H. 1981. Perceptual dimension of openness in vowels, Journal of the Acoustical Society of America, vol.69, pp. 1465-1475.

[4] Pape D., Mooshammer C., Fuchs S. and Hoole, P. 2005. Intrinsic pitch differences between German vowels /i:/, /I/ and /y:/ in a cross-linguistic perception experiment, *Proceedings of the PSP2005*, London, UK.

[5] Pape, D. and Mooshammer, C. 2006. Is Intrinsic pitch language-dependent? Evidence from a cross-linguistic vowel pitch perception experiment, *Proceedings of the MULTILING*, Stellenbosch, South Africa.

[6] Kishon-Rabin, L., Amir, O., Vexler, Y. and Zaltz Y. 2001. Pitch discrimination: Are professional musicians better than non-musicians? *J Basic Clin Physiol Pharmacol*. vol.12, no.2(Suppl.), pp.125-43.

[7] Gili Fivela B. and Zmarich C., 2005. Italian geminate under speech rate and focalization changes: kinematic, acoustic and perception data, *Proceedings of INTERSPEECH'2005-EUROSPEECH*, 2897-2900, Lisboa, Portugal.