Travaux du 19ème CIL | 19th ICL papers
Congrès International des Linguistes, Genève 20-27 Juillet 2013
International Congress of Linguists, Geneva 20-27 July 2013

19th International
Congress of Linguists
July 21-27 2013
Geneva - Switzerland

Martina URBANI

University of Padova
martina.urbani@unipd.it

The role of pitch range in L2 English by Italian speakers

poster presentation in session: 4 Phonology and Morphology (Marc van Oostendorp)

Published and distributed by: Département de Linguistique de l'Université de Genève, Rue de Candolle 2, CH-1205 Genève, Switzerland
Editor: Département de Linguistique de l'Université de Genève, Switzerland
ISBN:978-2-8399-1580-9

THE ROLE OF PITCH RANGE IN L2 ENGLISH BY ITALIAN SPEAKERS

Martina Urbani, University of Padova

martina.urbani@unipd.it

This study presents results on the differences of pitch range in selected utterances produced by American English native speakers and Italian learners of English. The hypothesis being tested is that Italian learners of English are influenced by their L1, thus transferring pitch range variation in their L2. The English sentences produced by Italians are expected to have overall higher pitch levels and narrower pitch span than those produced by Americans. To test this hypothesis, a cross-linguistic study was conducted by comparing pitch level and span in 15 sentences in English pronounced by 10 American English speakers from California and 10 Italian speakers from the North East of Italy. The corpus created consisted of 300 utterances (10 speakers x 15 sentences x 2 language groups). Cross-linguistic differences of pitch level and span were found across patterns used in different sentence types (yes-no questions vs. wh-questions vs. statements), with wh-questions and statements registering more significant differences than yes-no questions.

1. Introduction

What are the effects of pitch range in the perception of L2 speech? A number of studies investigated the nature of F0 span and level in cross-linguistic comparisons. However, only few experiments have focused on the real necessity to work on the elaboration of a general-agreed methodology. Some factors determining the existence of measurable and reliable differences in pitch values among speakers are the language spoken, age, body size, gender, socio-cultural background, regional accents, speech task, sentence type, and scales of measurement (Jenkins, 2000; Daly and Warren, 2001; Chun, 2002; Wells, 2006; Pickering, 2004). Documenting cross-linguistic differences in pitch range has proven to be difficult, because pitch is subjected to a wide range of inter-speaker and within-speaker variability, with data of speakers from different languages often overlapping. Studies in L2 intonation have shown that L2 learners tend to use the standard pitch range of their native language in their L2 (Chen, 2009; 2011). What is more, L2 language learners have consistently been reported as speaking with a narrower pitch range and less pitch variation than L1 language speakers (Mennen, 2006; 2007).

The comparison of speech material produced in English and Italian may present difficulties because of the inherent differences in the two languages. These differences regard both segmental

features (such as prevalence of voiced sounds over voiceless sounds in Italian, as compared to English) or prosodic factors (such as different numbers of syllables, placement of stress etc.). The analysis of pitch range is intricate and captivating for at least three reasons.

First, pitch range has been erroneously considered as a single unitary measure, while, actually, it is the result of two different dimensions: level and span (Mennen et al., 2012). While pitch level is a sort of reference line calculated over the rises and falls within each intonation contour, pitch span is a measure of the distance between the highest and the lowest F0 value in the contour (Ladd, 1996; Cruttenden, 1997; Gussenhoven, 2005). Hence, the description of pitch range variation is given by the sum of the changes in pitch level and span.

Second, previous research on pitch range suggested that the perception and the production of pitch variation have proper and distinctive features (for an overview see 't Hart et al., 1990). Thus, the non-trivial relation between perception and production should be treated seriously when analyzing the characteristics of pitch range across languages and populations. Some acoustic features of pitch range may be relevant in a study on production data and totally irrelevant in an study on perception data. Changes in the modulation of pitch patterns may be linguistically or communicatively relevant, depending on the kind of measures investigated. For example, purely acoustic studies can show dramatic changes of pitch range that may not be noticed and perceived by listeners. Thus, 'some apparently major pitch event may play a negligible role in perception while, conversely, a seemingly minor phonetic detail may prove indispensible' (Haan, 2002: 24).

Third, pitch range has a substantial impact on distinguishing native speakers from non-native speakers of different languages. Thus, it might also have a role in the perception of foreign accent. However, unlike other prosodic cues such as stress, rhythm, and speech rate, it is not clear to what extent pitch range may contribute to the detection of accented speech produced by L2 speakers. What is evident is that the speakers of different languages seem to have distinctive characteristics of pitch range (Gussenhoven, 2002; Chen et al., 2004; Mennen et al., 2012).

The present study is aimed at comparing the pitch range of Italians and Americans in order to provide evidence for cross-linguistic differences between L1 and L2 speakers. The focus of the analysis is based on the double nature of pitch range (i.e. pitch level and span) that is examined with an experimental approach relying on production data. Based on the idea that pitch range varies to different extent depending on the pitch contours of sentences, pitch range is analyzed in distinct sentence types such as yes-no questions, wh-questions and statements.

2. Research Questions

The aim of the present experiment is to find out whether or not pitch range considerably varies depending on sentence type and, most importantly, whether or not the English sentences produced by Americans have a pitch range similar to that of sentences produced by Italians.

The first hypothesis being tested has to do with the assumption that different intonation contours correspond to different pitch range. Pitch range (level vs. span) was measured for each utterance separately. Then, values were grouped depending on sentence type. Since it is generally agreed upon that yes-no questions (henceforth YNQ), wh-questions (henceforth WHQ), and statements (henceforth STM) are uttered with typical intonation contours, the dominant patterns for each sentence type may have an influence on pitch variation. Indeed, it has been shown that YNQ, WHQ, and STM are likely to have a different pitch range variation (Busà and Urbani, 2011).

The second hypothesis being tested is that Italian learners of English are influenced by their L1, thus transferring the L1 pitch range variation into their L2. On the basis of the results obtained in Urbani (2012), the English sentences uttered by the Italians are expected to have a narrower pitch span and higher pitch level than those produced by the Americans. Thus, L2 English sentences might have a slightly monotonous tone and rather flat pitch contour, due to the prosodic transfer from L1. If this hypothesis is confirmed, the pitch range shown by the Italians in their L2 speech will be much less dynamic and varied than the pitch range shown by the Americans in their L1 speech.

The null hypotheses being tested are (i), the F0 level and span values are the same across sentence types (YNQ vs. WHQ vs. STM), (ii) the F0 measures for F0max, F0mean, F0min, and STrange are the same in the two native language groups (Americans vs. Italians).

3. Subjects

Ten female adult native speakers of American English and ten female adult native speakers of Italian volunteered for the study. All the American participants were speakers of American English, they came from California and were students at the University of California – Los Angeles. All the American subjects were also proficient in Italian at different levels and they had lived in Italy for up to one year, by taking part into exchange programs held in Padua, Bologna, Florence and Milan. All the American participants were university students in different departments at UCLA: Italian, as a major or minor (6 students); Linguistics (1 student); History (1 student); Psychology (1 student); Marine Science (1 student). Eight students were undergraduate while two students were doing a Master degree. The age of the participants ranged from 20 to 26 years (mean age: 21,5 years). None

of the speakers reported any speech, hearing or communication disorder and they were all non-smokers. There was no screening for formal training in music or singing.

All the Italian speakers were either graduate students at the University of California – Los Angeles or graduate students at the Università degli Studi di Padova. They were university students in different departments at the University of Padova and UCLA: English literature and/or linguistics (6 students); Linguistics (2 students); Italian (1 student); French (1 student). All the Italian subjects but one were graduate students; two students were doing a Master degree while seven students were pursuing a Ph.D. The age of the participants ranged from 22 to 31 years (mean age: 27,8 years). Other than age, homogeneity of the subjects was controlled for the Italian variety they spoke: all the participants in the experiment were speakers of the Northern-East Italian variety (Veneto area).

Subjects were asked to rate how often they watched movies or TV programs broadcasted in their L2: 75% of Italians answered that they watch English programs at least once a week, only 20% of Americans regularly watch Italian programs. When inspecting the information about competence in L2 (years of learning, proficiency level, and time spent abroad), a great homogeneity in the two groups was found. In particular, all the American subjects were proficient in Italian. They were students of Italian (mostly at the intermediate level) and spent several months in Italy to practice their language skills (always periods inferior to one year). All the Italian subjects were proficient in English. They were students of English (mostly at the advanced level) and spent several months in the United States to practice their language skills. Eight students had spent less than one year in the United States while two students had been living for more than one year in the United States, at the time of the recording.

In sum, the total number of subjects selected for the experiment was 20: 10 American subjects and 10 Italian subjects. They were all females, in the same age and with similar competence in English and Italian. The author personally knows all the participants she recruited for the experiment. They gave their consent for the treatment of their personal data and volunteered for the experiment without receiving any monetary compensation.

4. Materials

This study compares native and non-native productions of 15 English sentences produced by 10 American and 10 Italian subjects. In the data set, all the sentences were divided into three groups depending on sentence type: 5 YNQ, 5 WHQ, and 5 STM. Every sentence was read by each participant at least twice (when the subjects were misreading a sentence, more repetitions were

necessary). Only two repetitions were retained for every sentence. In table 1, the 15 sentences created for the analysis of pitch range are shown:

	English sentences
Yes/no questions	Do you need any money? Have we met before? Are you still there? Can you open the door? Do you wanna come for dinner?
Wh-questions	Where were you when the money ran out? Why are you selling meat? What was her name again? What are you doing there? What's wrong with you?
Statements	Now you are going away. I hope I can see you on Monday. We should go and visit your uncle. I know you are leaving today. You should go to Hawaii.

Table 1. English sentences for the analysis of pitch range grouped according to sentence type: yes/no questions, wh-questions and statements.

The materials created for the present experiment had to conform to specific standards. Sentences had to be short and they had to contain easy to pronounce words. Voiced sounds were prevalent over voiceless sounds because the program used for the analysis fails to capture the pitch track of voiceless sounds. In particular, the three groups of sentences had the prosodic characteristics of different sentence types: YNQ vs. WHQ vs. STM. The corpus created consisted of 300 utterances (5 sentences x 20 speakers x 3 sentence types).

5. Procedure

The subjects were asked to read aloud short sentences in a natural way. The text was read aloud by 10 American English female speakers from California and 10 Italian female speakers from the North East of Italy (Veneto area). Both the American English and the Italian subjects read the materials in English. Data were extracted from two different groups: (1) Americans speaking English, (2) Italians speaking English.

The materials were collected by the author in separate sessions, in a three weeks period at the Linguistics Department of the University of California – Los Angeles. The audio files were recorded and digitally acquired in a sound-attenuated booth in the UCLA Phonetic Laboratory. They were collected using a Shure SM10 head-mounted microphone, recorded direct-to-disk on another computer located outside the sound booth, and digitized at a sample frequency of 44.1 KHz and a 32 bit quantization rate, using an AudioBox. By using a sampling rate of 44.1 KHz (i.e. CD quality), it was possible to collect data with excellent quality. After recording the short sentences, the author saved the data and labeled them as separate WAV audio files with Praat (Boersma and Weenick, 2010).

During the recording session, subjects were instructed to read sentences with a natural conversational intonation. No indication was given about the intonation they had to use in the different types of sentence. The subjects were constantly monitored by the author while they were reading the sentences and they were required to repeat any sentence when they misread it. In some cases, when the speakers did not feel comfortable with the utterance pronounced, they asked to do the recording again. Before starting with the recording process, subjects were permitted to read silently the sentences in order to familiarize themselves with them. Each recording session lasted about 15 minutes. At the end of the session, every subject was asked to fill in a questionnaire. Speakers were requested to indicate their first and last name; age; birth place; sex; native language; second languages and proficiency levels; university status; periods abroad; ways of learning a second language and its daily use.

6. Method

The data were analyzed by following the method proposed by Mennen et al. (2012). Thereby pitch values were analyzed and compared across groups by calculating long-term distributional (LTD) and linguistic measures.

Linguistic measures were calculated by manually annotating every sentence. The beginning and the end of every sentence were marked with a I for the initial pitch and a FL for a final low pitch or a FH for a final high pitch. Local peaks on prominent syllables were marked as H* while local peaks on non-prominent syllables were marked as H. When peaks were placed at the phrase starting point they were labeled with an additional ‘i’ to signal their initial position, e.g. H*i and Hi. Valleys never appeared in initial position and they were labeled according to their prominence

status. Local valleys on prominent syllables were identified by the label L* while local valleys on non-prominent syllables were identified by the label L.

LTD measures were based on the analysis of F0 distribution. Values of F0 maximum (F0max), F0 minimum (F0min), F0 mean (F0mean) and F0 median (F0median) were calculated over the entire sentences to measure pitch level. Measures analyzed for pitch span were: F0 maximum minus F0 minimum (max-min F0) in Hz and ST. Values for LTD measures were obtained automatically by inquiring pitch information in Praat such as minimum, maximum, range, average, and standard deviation. The same protocol was used to calculate all measures using the same standards and procedure. Pitch tracking was performed with a standard algorithm based on the autocorrelation method. This algorithm is the standard option to process speech and detect pitch locations in Praat.

7. Results

The following sections present an analysis of 300 sentences that are divided into three sentence types: yes-no questions, wh-questions, and statements (100 YNQ, 100 WHQ, and 100 STM). Thus, the present study provides an overview of how pitch range variation is associated to specific sentence types in English as an L1 and L2. The results were obtained from the elaboration of data on linguistic and LTD measures. By systematically comparing pitch range of YNQ, WHQ and STM, a three-way contrast in their F0 realizations is described.

7.1 Linguistic measures

After placing linguistic landmarks at peaks and valleys, a Praat script created by Mennen at al. (2012) was used to calculate the F0 of every pitch point. The values were averaged across speakers and language groups (the Americans speaking English vs. the Italians speaking English). The values of linguistic measures were calculated for pitch level (tab. 2) and pitch span (tab. 3). In table 2, measures calculated for level were grouped depending on the native language of the speakers and the sentence types.

For F0 level, L*, L and FL were counted as the measures of valleys, that is, the bottom line of the pitch contour. H*i, H*, Hi and H identified peaks within the intonation contour, that is, the top line of the pitch contour.

	YNQ		WHQ		STM	
Linguistic level	*AmE*	*It*	*AmE*	*It*	*AmE*	*It*
I	271	189	241	267	207	227
Hi	318	301	332	318	299	272
H*i	278	266	358	346	410	350
H*	265	269	294	315	378	277
H	159	193	265	332	271	248
L*	133	177	146	248	155	192
L	182	229	132	233	139	233
FH	416	353	_	_	_	_
FL	_	_	114	253	126	187

Table 2. Overview of linguistic measures for level. Mean values for each landmark were calculated in Hz for the American English (AmE) and the Italian (It) female subjects. Data were divided for each label into three groups depending on the sentence types: YNQ, WHQ and STM.

The sentence initial target point, I, and final target points, FH and FL, were included because they stand for reference points for the F0 movements across the contours. As shown in table 2, the landmark FL was not included in the measures of the YNQ because YNQ are characterized by intonation patterns ending in final rises (FH) and not in final falls (FL). By contrast, the WHQ and the STM did not show measures for the landmark FH because WHQ and STM are characterized by intonation patterns ending in final falls (FL) and not in final rises (FH).

In the YNQ, the highest F0 values were reached at the FH target point (i.e. 416 Hz for the Americans and 353 Hz for the Italians). The lowest F0 values were reached at the L* target point (i.e. 133 Hz for the Americans and 177 Hz for the Italians). In the WHQ, the highest F0 values were reached at the H*i (i.e. 358 Hz for the Americans and 346 Hz for the Italians). The lowest F0 values were reached at the FL target point for the Americans (i.e. 114 Hz) and at the L target point for the Italians (i.e. 233 Hz). In the STM, the highest F0 values were reached at the H*i target point (i.e. 410 Hz for the Americans and 350 Hz for the Italians). The lowest F0 values were reached at the FL target point (i.e. 126 Hz for the Americans and 187 Hz for the Italians).

For F0 span, targets points were calculated to describe the pitch movements along the measures: I-L*, H*i-L*, H*-L*, FH-L*, I-FL, H*-FL, H*i-FL. As shown in table 3, landmarks such as Hi, H and L were not included in the measures for span because their values were less extreme than those of H*I, H* and L*.

	YNQ		WHQ		STM	
Linguistic span	*AmE*	*It*	*AmE*	*It*	*AmE*	*It*
I – L*	138	12	95	19	52	35
H*i – L*	145	89	212	98	255	158
H* – L*	132	92	148	67	223	85
FH – L*	283	176	_	_	_	_
I – FL	_	_	127	14	81	40
H* – FL	_	_	180	62	252	90
H*i – FL	_	_	244	93	284	163

Table 3. Overview of linguistic measures for span. Mean values for each landmark were calculated in Hz for the American English (AmE) and the Italian (It) female subjects. Data were divided for each label into three groups depending on the sentence types: YNQ, WHQ and STM.

In the YNQ, results for span show that the widest pitch excursions were reached by the FH-L* measure (i.e. 283 Hz for the Americans and 176 Hz for the Italians), while the narrowest span values were obtained by the H*-L* measure for the Americans (i.e. 132 Hz) and by the I-L* measure for the Italians (i.e. 12 Hz). In the WHQ, results for span show that the widest pitch excursions were reached by the H*i-FL measure for the Americans (i.e. 244 Hz) and by the H*i-L* measure for the Italians (i.e. 98 Hz). The narrowest span values were obtained by the I-L* measure for the Americans (i.e. 95 Hz) and by the I-FL measure for the Italians (i.e. 14 Hz). In the STM, results for span show that the widest pitch excursions were reached by the H*i-FL measure (i.e. 284 Hz for the Americans and 163 Hz for the Italians), while the narrowest span values were obtained by the I-L* measure (i.e. 52 Hz for the Americans and 35 Hz for the Italians).

The figures 1-3 show the linguistic measures calculated for the pitch patterns of YNQ, WHQ, and STM. The values obtained for the Americans are plotted against the values obtained for the Italians. Data on the linguistic measures were averaged and displayed on a graph containing two patterns: the Americans speaking English (the blue line of diamonds), the Italians speaking English (the red line of squares).

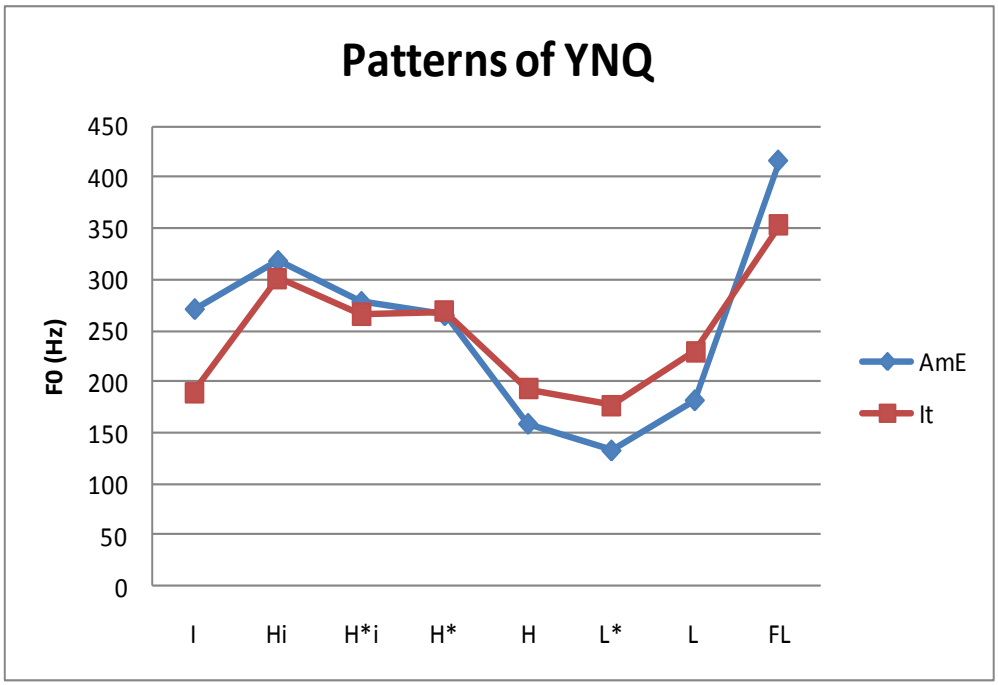


Figure 1. Linguistic measures calculated for the American (red squares) and Italian (blue diamonds) female speakers. Measures are described along YNQ patterns signaled by pitch points corresponding to linguistic landmarks.

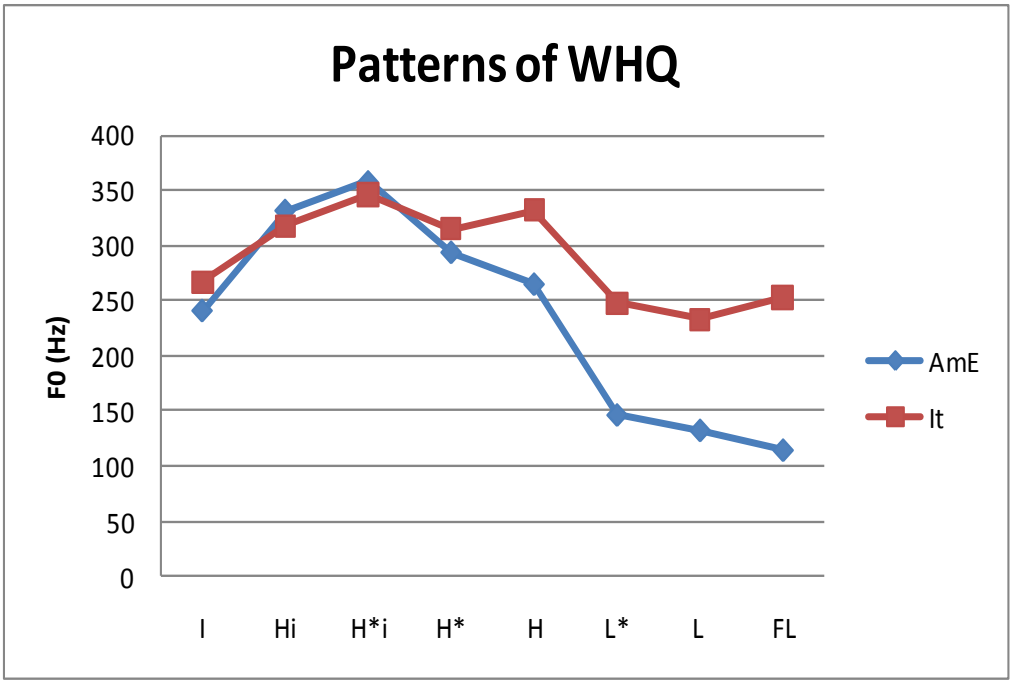


Figure 2. Linguistic measures calculated for the American (red squares) and Italian (blue diamonds) female speakers. Measures are described along WHQ patterns signaled by pitch points corresponding to linguistic landmarks.

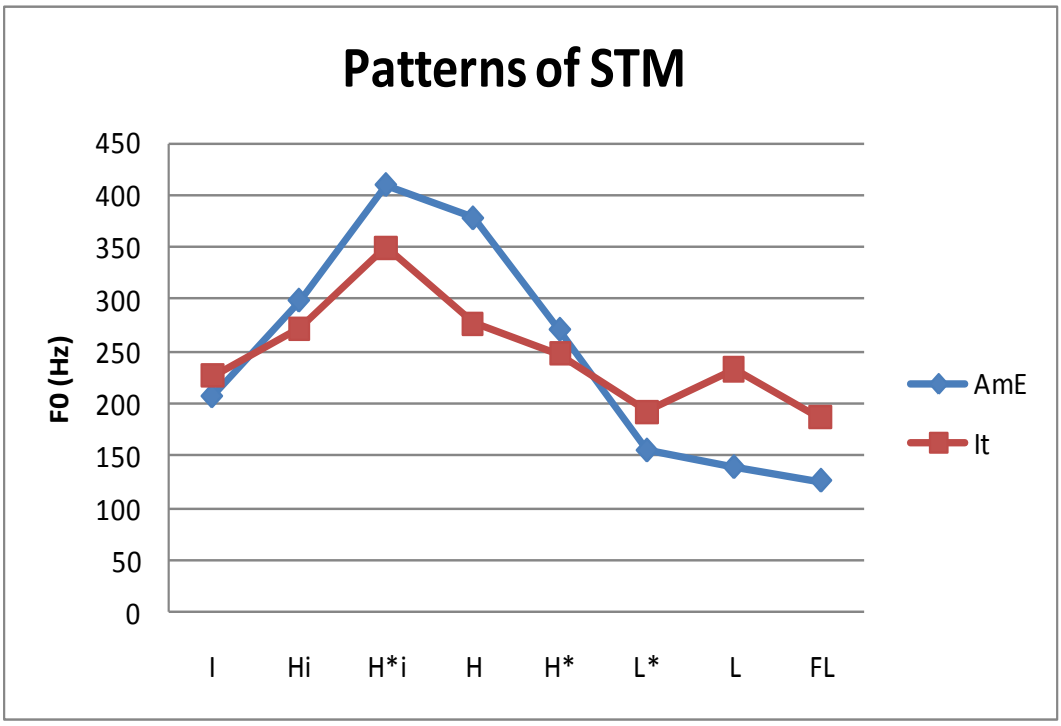


Figure 3. Linguistic measures calculated for the American (red squares) and Italian (blue diamonds) female speakers. Measures are described along STM patterns signaled by pitch points corresponding to linguistic landmarks.

The **patterns for YNQ** (fig. 1) evidenced a clear similarity between the pitch contours of the Americans and the Italians, despite the fact that the Americans had higher values than the Italians at the beginning and at the end of the patterns: compare the values obtained for the initial pitch point I (i.e. 271 Hz for the Americans and 189 Hz for the Italians) and for the final rise FH (i.e. 416 Hz for the Americans and 353 Hz for the Italians). In addition, the mean difference across the H*i, Hi, H*, H, L*, L measures calculated for the Americans and the Italians was 26,33 Hz. By contrast, the mean difference between the Americans' and the Italians' values calculated for the I and the FH measures was 82 Hz for I and 63 Hz for FH. This means that, unlike the I and the FH measures, most of the pitch points measured in the YNQ obtained similar values for the Americans and the Italians.

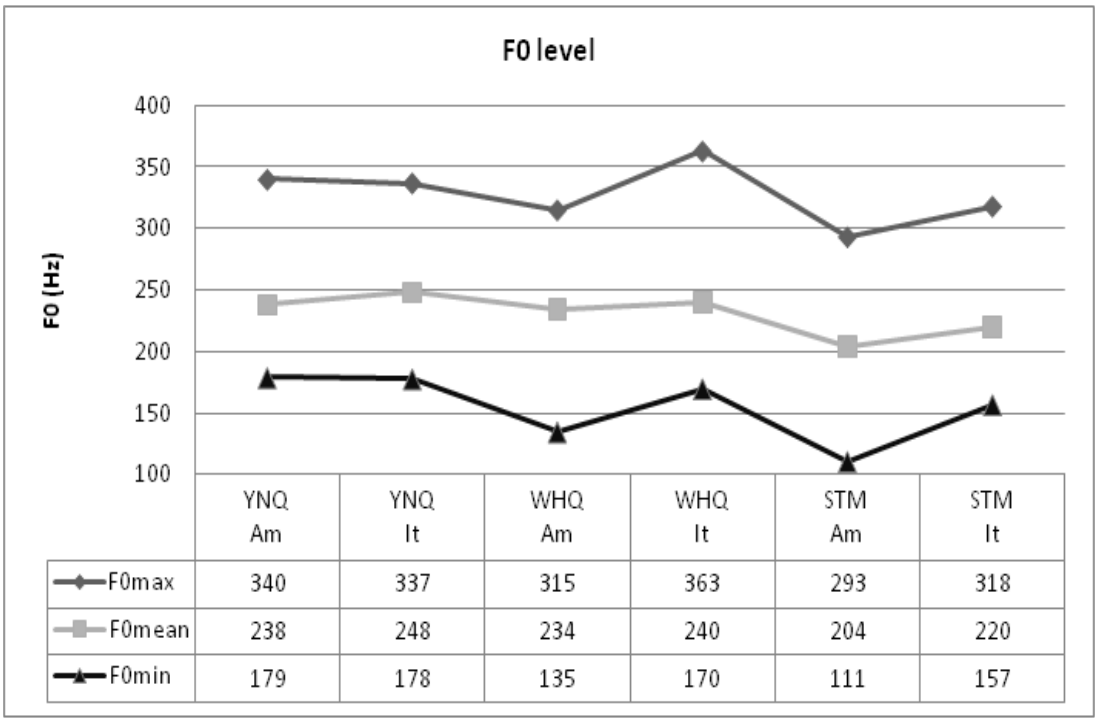
The **patterns for WHQ** (fig. 2) showed fairly similar values obtained for initial and high peaks (I, H*i, Hi, H*) and different values obtained for final falls and valleys (L*, L, FL). The non-initial and non-prominent peak H had a considerably lower value in the pitch pattern used by the Americans than the Italians (i.e. the value for H was 265 Hz for the Americans and 332 Hz for the Italians). In addition, the mean difference between the Americans' and that Italians' values calculated for initial and high peaks (I, H*i, Hi, H*) was 18,25 Hz. By contrast, the mean difference

in final falls and valleys (L*, L, FL) measures calculated for the Americans and the Italians was 114 Hz. This means that, the WHQ patterns were characterized by similar values for initial and high peaks and different values for final falls and valleys. Unlike the Italians, the Americans seem to realize a sharp slope at the end of the WHQ by reaching the lowest values of their pitch patterns.

The **patterns for STM** (fig. 3) showed fairly similar values obtained for some initial and non-prominent high peaks (I, Hi, H) and different values obtained for accented peaks (H*, I and H*) and final lows (L, FL). The Americans had higher values than the Italians at the peaks and at the valleys of the utterances. Extremely different values for the Americans and the Italians were obtained for the initial peak H* (i.e. 378 Hz for the Americans and 277 Hz for the Italians) and for the final fall FL (i.e. 126 Hz for the Americans and 187 Hz for the Italians). In addition, the mean difference between the Americans' and the Italians' values calculated in the I, Hi, H, L* measures was 26,75 Hz. The mean difference in the H*i and the FL measures calculated for the Americans and the Italians was 60 Hz for H*i and 61 Hz for FL. The highest differences between the pitch values obtained for the Americans and the Italians were reached for the H* measure (i.e. 101 Hz) and the L measure (94 Hz). This means that, the STM pitch line produced by the Italians was realized with a narrower pitch span than that produced by the Americans. Compared to the Americans' STM, the Italians' STM resulted as more flat and compressed. Unlike the Italians, the Americans realized much sharper rises and falls, by reaching the highest and lowest values in the STM patterns.

7.2 LTD measures

The processing of LTD measures required the analysis of a total of 300 utterances (5 sentences x 3 sentence types x 10 subjects x 2 native languages). Namely, every participants in the experiment, 10 Americans and 10 Italians, produced 30 sentences: 10 YNQ, 10 WHQ, and 10 STM. The graph in table 4 shows the mean values obtained for different measures of F0 level (i.e. F0max, F0mean, and F0min) for the two language groups (i.e. American native speakers and Italian learners of English), evidencing clear differences in pitch level patterns across sentence types (i.e. YNQ, WHQ, and STM).



	YNQ Am	YNQ It	WHQ Am	WHQ It	STM Am	STM It
F0max	340	337	315	363	293	318
F0mean	238	248	234	240	204	220
F0min	179	178	135	170	111	157

Table 4. F0 max, mean and min values in Hz by American and Italian subjects obtained in three sentence types: YNQ, WHQ, and STM.

F0 level considerably varied across sentence types and language groups, with WHQ and STM showing more significantly differences between the productions of the Americans and the Italians. It is clear, from the graph in table 4, that F0 level for WHQ and STM was shifted downwards in the utterances produced by the Americans, as compared to F0 level in the Italian utterances. Surprisingly, YNQ showed very similar results for the Americans and the Italians. In the sentences produced by the Americans, every F0 level measure (i.e. F0max, F0mean, and F0min) reached the highest values in YNQ and the lowest in STM, with values for WHQ in between. Also the sentences produced by the Italians were in line with this trend, with the exception of WHQ. In fact, F0max was higher for WHQ (363 Hz) than YNQ (337 Hz). Since, WHQ are commonly uttered with falling contours, one would expect that they have rather low F0max values. Nevertheless, this didn't happen.

F0 span was calculated in ST because logarithmic scales (e.g. ST) better than linear scales (e.g. Hz) manage to capture the excursions between F0 values (Daly and Warren, 2001; Nolan, 2003; Mennen et al. 2012). The bars in table 5 show the distribution of F0 values in ST, for the three sentence groups (YNQ vs. WHQ vs. STM), showing evident differences in pitch span patterns across the language groups (Americans vs. Italians).

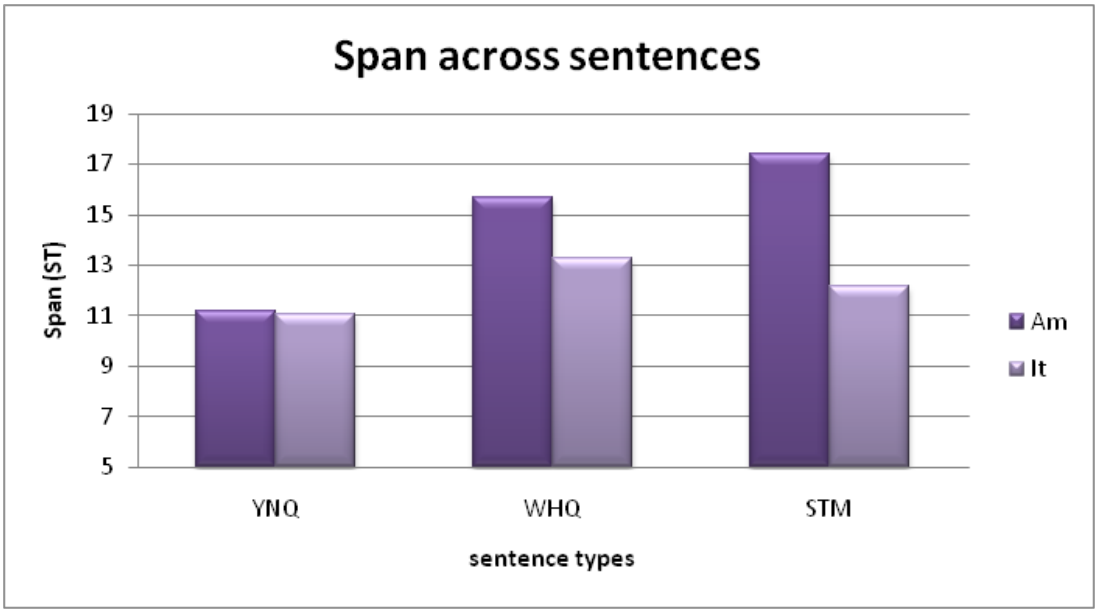


Table 5. Span values across sentences (YNQ, WHQ, and STM) by the American and the Italian subjects.

YNQ reached fairly similar pitch span values, with span in American YNQ (11.19 ST) slightly larger than in Italian YNQ (11.03 ST). Both WHQ and STM showed great differences in the span values of the American vs. Italian utterances. Span values for WHQ were 15,65 ST for the Americans and 13,29 ST for the Italians; span values for STM were 17,42 ST for the Americans and 12,19 ST for the Italians. In sum, the Americans' span was larger than that of the Italians in every sentence type.

A paired t-test assuming equal variances between groups was selected as the preferred method to analyze data in this experiment, mostly for two reasons: 1) it is one of the most commonly used tests in phonetics studies (Rasinger, 2008; Lane, 2012); 2) it suits the kind of variables investigated in this experiment. The F0 values were loaded into excel to obtain inferential statistics.

For the **factor *'sentence type'*** (YNQ vs. WHQ vs. STM), two-sample paired t-tests were run for the dependent variables F0 level and span, calculated for the sentences produced by the Americans and the Italians. Statistical significance for the 'sentence type' factor was measured in a total of 12 t-tests. The dependent variables were two measures for pitch range (F0 level and span), the independent variables were the native languages of the speakers (American vs. Italian). Statistically significant differences across sentence types (YNQ vs. WHQ vs. STM) were largely proved by several t-tests showing that both Italians and Americans modify F0 level and span, depending on sentence type.

For the **factor *'native language'***, two-tail paired t-tests were run for the dependent variables (F0max, F0mean, F0min, and STrange), separately calculated for every sentence type. Statistical significance for the 'native language' factor was measured in a total of 12 t-tests with F0 measures (F0max, F0mean, F0min, and STrange) as dependent variables and sentence type (YNQ vs. WHQ vs. STM) as independent variables. As shown in figures 4-7, the two-tail paired t-tests show that there was a significant effect for the 'native language' factor across sentences produced by American native speakers and Italians native speakers.

Table 6 shows average F0mean, F0max, F0min and STrange measures calculated in YNQ, WHQ and STM. Values are compared across American and Italian subjects to test whether or not differences are significant.

Measure	Sentence	AM	IT	p-value	significance
F0 mean	YNQ	238 Hz	248 Hz	.032	*
	WHQ	234 Hz	240 Hz	.094	
	STM	204 Hz	220 Hz	< .001	*
F0max	YNQ	340 Hz	337 Hz	.635	
	WHQ	315 Hz	363 Hz	< .001	*
	STM	293 Hz	318 Hz	< .001	*
F0min	YNQ	179 Hz	178 Hz	.933	
	WHQ	135 Hz	170 Hz	< .001	*
	STM	111 Hz	157 Hz	< .001	*
STrange	YNQ	11.19 ST	11.03 ST	.714	
	WHQ	15.65 ST	13.26 ST	< .001	*
	STM	17.42 ST	12.19 ST	< .001	*

Table 6. Statistics and effect size for t-tests. An asterisk denotes significance after Bonferroni correction.

T-tests showed that the between-subject factors 'native language' and 'sentence type' reached significance for most dependent variables, characterizing F0 level and span. This suggests that the native language of the subjects (L1) plays a relevant role in pitch variation. As far as F0 level is concerned, the results showed that F0 level considerably varied across sentence types and L1 speakers (i.e. the Americans and the Italians), with YNQ and STM showing more significant differences between the productions of the American and the Italians. The F0 level for WHQ and STM was

shifted downwards in the utterances produced by the Americans, as compared to the Italian utterances. F0max and F0min were significantly higher for the Italians than for the Americans in WHQ and STM. By contrast, YNQ obtained very similar results in F0max and F0min values for the Americans and the Italians.

As far as F0 span is concerned, data for WHQ and STM evidenced neat differences across language groups. No significant difference was found for F0 span in YNQ produced by the Americans and the Italians. Across sentences, F0 span was significantly higher for the Italians than for the Americans, with the exception of YNQ which displayed almost identical F0 span values for the Americans and the Italians.

8. Conclusion

This study was designed to determine how pitch level and span vary in English sentences produced by American and Italian females. In particular, pitch variation was tested across YNQ, WHQ and STM in order to identify a specific model of pitch patterns for each sentence type produced in English as L1 and L2. The aim of the present experiment was to find out whether or not pitch range considerably varies depending on the sentence types (e.g. YNQ vs. WHQ vs. STM) and, most importantly, whether or not English sentences produced by the Americans have a pitch range similar to that of sentences produced by the Italians.

Five dependent variables (F0mean, F0median, F0max, F0min, and span) were tested. The results showed that Americans' F0 span was larger than that of the Italians while the Americans' F0 level was lower than that of the Italians. This gives indication about the fact that pitch range is sensibly different in English sentences uttered by native speakers (in this case, Americans) and non-native speakers (in this case, Italians). Data also proved that the Italian subjects produce all sentence types with a narrower pitch span and a higher pitch level than the Americans. This means that English as L2 is more high-pitched than English as L1. This finding is in line with data from other languages, such as Swedish, standard Chinese, Japanese, and Hungarian (Bolinger, 1972, 1978; Ohala, 1983; van Beezoijen, 1995; Gussenhoven, 2002; Yuasa, 2008). A considerable drop in pitch span was observed for the sentences produced by the Italians speaking English. This is an effect of one main factor: the English L2 pitch span is narrower than the English L1 pitch span.

Important differences were registered across sentence types. YNQ were uttered with similar pitch level and span by both the Americans and Italians. High pitch level and wide pitch span are common traits in YNQ across languages. Consequently, the Italian learners of English may either be

familiar with the English final rising patterns or may simply transfer patterns from their L1 to their L2. The Italian subjects of this study, whether subconsciously or consciously, successfully managed to approach the American model for YNQ pitch contours. By contrast, the American and Italian subjects differed in their production of WHQ and STM pitch contours. The data showed that the F0 level for WHQ and STM was shifted upwards in the acoustic space of the utterances produced by the Italians, as compared to those produced by the Americans. Also the data calculated for pitch span evidenced neat differences across WHQ and STM produced by the Americans and the Italians. Generally, the dependent variables measured for pitch level (F0mean, F0median, F0max, F0min) were significantly higher in the utterances produced by the Italians. On the other hand, the dependent variable of pitch span (STrange) was significantly wider in the utterances produced by the Americans.

The results of the present study lead to the conclusion that F0 span has a significant role in distinguishing pitch patterns displayed in English as L1 and L2. The pitch span of the Italians speaking English is considerably narrower than that of the American English native speakers. What is more, the experimental data gathered across sentence types indicate that the mode of sentences is better captured by the measures of F0 span than level.

References

Boersma, P. and Weenink, D. (2010). *Praat: doing phonetics by computer* [computer software, version 5.3.32]. Amsterdam: University of Amsterdam. Available and downloadable at: http://www.praat.org

Bolinger, D. (1972). *Intonation: selected readings*. Harmondsworth: Penguin.

Bolinger, D. (1978). Intonation across languages. In: Greeberg, J.H., Ferguson, C.A., Moravcik, E.A. (eds.), *Universals of Human Language*, 2. Stanford, CA: Stanford University Press, 471-524.

Busà, M.G. and Urbani, M. (2011). A cross linguistic analysis of pitch range in English L1 and L2. *Proc. 17th International Conference of Phonetic Sciences (ICPhS)*, Hong Kong, 380-383.

Chen, A., Gussenhoven, C., Rietveld, T. (2004). Language specificity in perception of paralinguistic intonational meaning. *Language and Speech*, 47 (4), 311-349.

Chen, A. (2009). Perception of paralinguistic intonational meaning in a second language. *Language Learning*, 59, 2, 448-451.

Chen, A. (2011). What's in a rise. Evidence for an off-ramp analysis of Dutch intonation. *Proc. 17th International Conference of Phonetic Sciences (ICPhS)*, Hong Kong, 380-383.

Chun, D.M. (2002). *Discourse intonation in L2. From theory and research to practice*. Amsterdam: John Benjamins.

Cruttenden, A. (1997). *Intonation* (2nd ed.). Cambridge: Cambridge University Press.

Daly, N. and Warren, P. (2001). Pitching it differently in New Zealand English: Speaker sex and intonation patterns. *Journal of Sociolinguistics*, 5 (1), 85-96.

Gussenhoven, C. (2002). Intonation and interpretation: phonetics and phonology. *Proc. Speech Prosody*, Aix-en-Provence, 47-57.

Gussenhoven, C. (2005). Transcription of Dutch Intonation. In: Sun, A.-J. (ed.), *Prosodic Typology: The Phonology of Intonation and Phrasing*. Oxford: Oxford University Press.

Haan, J. (2002). *Speaking of questions. An exploration of Dutch question intonation*. LOT dissertation series 52, LOT: Utrecht.

Jenkins, J. (2000). *The phonology of English as an international language*. Oxford: Oxford University Press.

Ladd, D.R. (1996). *Intonational Phonology*. Cambridge: Cambridge University Press.

Lane, D.M., (eds.) (2012). Online Statistics Education: A Multimedia Course of Study. Available at: http://onlinestatbook.com/Online_Statistics_ Education.pdf (Last accessed 14th Oct. 2014).

Mennen, I. (2006). Phonological and phonetic influences in non-native intonation. In: Trouvain, J. and Gut, U. (eds.), *Non-native Prosody: Phonetic Descriptions and Teaching Practice*. Berlin: Mouton De Gruyter, 1-18.

Mennen, I., Schaeffler, F., Docherty, G. (2007). Pitching it differently: a comparison of the pitch ranges of German and English speakers. *Proc. 16th International Conference of Phonetic Sciences (ICPhS)*, Saarbrücken, 1769-1772.

Mennen, I., Schaeffler, F., Docherty, G. (2012). Cross-language difference in f0 range: a comparative study of English and German. *Journal of the Acoustical Society of America*, 131 (3), 2249-2260.

Nolan, F. (2003) Intonational equivalence: an experimental evaluation of pitch scales. *Proc. 15th International Conference of Phonetic Sciences (ICPhS),* Barcelona, 771-774.

Ohala, J. J. (1983). Cross-language use of pitch: An ethological view. *Phonetica*, 40, 1-18.

Pickering, L. (2004). The structure and function of intonational paragraphs in native and nonnative speaker instructional discourse. *English for Specific Purposes*, *23*, 19-43.

Rasinger, Sebastian M. (2008). *Quantitative Research in Linguistics*. London: Continuum.

't Hart, J., Collier, R., Cohen, A. (1990). *A perceptual study of intonation*. Cambridge: Cambridge University Press.

Urbani, M. (2012). Pitch range in L1/L2 English. An analysis of F0 using LTD and linguistic measures. In: Busà, M.G. and Stella, A. (eds.). *Methodological Perspectives on Second Language Prosody*. *Papers from ML2P 2012*. Padova: Cleup, 79-83.

van Bezooijen, R. (1995). Sociocultural aspects of pitch differences between Japanese and Dutch women. *Language and Speech*, 38, 253-265.

Wells, J.C. (2006). *English Intonation*. *An Introduction*. Cambridge: Cambridge University Press.

Yuasa, P.I. (2008). *Culture and Gender of Voice Pitch*. *A Sociophonetic Comparison of the Japanese and Americans*. London: Equinox.