DISENTANGLING STRESS FROM ACCENT IN SPANISH
PRODUCTION PATTERNS OF THE STRESS CONTRAST IN DEACCENTED SYLLABLES*

MARTA ORTEGA-LLEBARIA[1] & PILAR PRIETO[2]
[1]*University of Texas at Austin,* [2]*ICREA & Universitat Autònoma de Barcelona*

Abstract

According to Sluijter and colleagues (1996b, 1997), stress is independent from accent because it has its own phonetic cues: stressed vowels are longer and have flatter spectral tilts than their unstressed counterparts. However, Campbell and Beckman (1997) show that, for American English, these duration and spectral tilt patterns are a consequence of vowel reduction: when unreduced vowels with different levels of stress (primary and secondary stress) are compared, duration and spectral tilt do not correlate with the stress difference. This paper contributes to the above discussion by examining the stress contrast in deaccented syllables in Spanish. Since Spanish has no phonological vowel reduction, it constitutes a good test case for the above hypotheses. Moreover, this study attempts to disentangle the correlates of stress from those of accent, something which has thus far not been done in the traditional literature on Spanish stress. The results indicate that stress contrast in Spanish is maintained in deaccented contexts by differences in duration, spectral tilt, and to a lesser extent, vowel quality.

1. *Introduction*

In this article we examine the phonetic characterization of the stress contrast in Spanish in accented and deaccented syllables. Stress (or 'primary stress') is a structural linguistic property of a word which specifies which syllable will be 'stronger', i.e. more prominent than the others. In stress-accent languages, stressed syllables serve as the landing site for accents, which are signalled acoustically by a pitch movement (Bolinger 1958, 1961; Pierrehumbert 1980; Beckman 1986; Ladd 1996; Beckman & Edwards 1994; Sluijter & van Heuven 1996a, 1996b, among others). However, not all

* The results of this experiment were presented at the 2005 PaPI Conference (Bellaterra, June 2005). We would like to thank the audience at this conference, and especially Lluïsa Astruc, Sónia Frota, Barbara Gili-Fivela, José Ignacio Hualde, Bob Ladd, Carme de la Mota, and Marina Vigário for very useful feedback. We are also grateful to Laura Colantoni and two anonymous reviewers for their comments on an earlier version of this paper. Finally, we thank the informants who kindly participated in the production experiment (Néstor, Ángel, Esther, Cristina, and Mari Carmen). This research has been funded by grants 2002XT-00032, 2001SGR 00150, and 2001SGR 00425 from the *Generalitat de Catalunya* and HUM2006-01758/FILO from the Ministry of Science and Technology of Spain.

syllables with primary stress are accented in all discourse contexts: the presence or absence of a pitch accent depends on the larger prosodic structure in which the lexical item is found. Thus, there exist at least three levels of syllabic prominence: unstressed, stressed and accented, and stressed but not accented.

Thus far, acoustic cues to stress prominence in Spanish have been studied in words and sentences spoken in intonation patterns that exhibited covariation between stress and accent. In other words, all stressed syllables also had a pitch accent, while unstressed syllable were deaccented (see Navarro Tomás 1914, 1964; Contreras 1963, 1964; Quilis 1971; Gili Gaya 1975; Solé 1984; Canellada & Kuhlman-Madsen 1987; Llisterri, Machuca, de la Mota, Riera & Ríos 2003, among others). As a consequence the cues to stress could not be distinguished from cues to accent in the results of these studies and, not surprisingly, the researchers found that pitch movements accompanied stressed syllables. Of these authors, only Navarro Tomás claimed that, in Spanish, the strongest cue to stress was a local increase in loudness or intensity and established the idea that stress in Spanish was mainly an 'intensity stress' (the so-called *acento de intensidad*), while relating pitch movements to intonation[1]. To our knowledge, the only results available on the production of stress cues in Spanish while controlling for the effects of accent are those of Ortega-Llebaria (2006). She finds evidence that supports Navarro-Tomas's hypothesis, namely that stress and accent in Spanish are related to different phonetic cues, i.e. pitch relates to accent while intensity cues stress. However, her study was limited to oxytone words.

In terms of other languages, Sluijter and colleagues' experiments on the correlates of stress in Dutch and American English were among the first that controlled for stress and accent covariation (Sluijter & van Heuven 1996b; Sluijter, van Heuven & Pacilly 1997). They found that stressed syllables were longer and had flatter spectral tilts than their unstressed counterparts, regardless of whether they bore a pitch accent or not. Thus, they too found that intensity cues related to stress, not accent, and concluded that stress was not a weaker degree of accent:

> One would expect to observe lower values along all measure correlates in stressed syllables of unaccented words. However, what we do observe is weakening along only those dimensions that are related to the omission of accent-lending pitch movements. (Sluijter & van Heuven 1996b:2483)

[1] "El acento de intensidad, que en estado actual de la pronunciación española influye más que ningún otro elemento en la estructura prosódica de nuestras palabras, proviene directamente, en la mayor parte de los casos, de la acentuación latina." (Navarro Tomás 1914:176, sec. 159)
"A veces, bajo una misma forma, se dan dos o tres palabras distintas, que fonéticamente sólo se diferencian por el lugar en que cada una de ellas corresponde al acento de intensidad: *límite-limite-limité*, *célebre-celebre-celebré*, *depósito-deposito-depositó*... (..) El oído español es evidentemente más sensible a las modificaciones de intensidad que a las de otros elementos fonéticos." (Navarro Tomás 1914:177, sec. 159)

Campbell and Beckman (1997) replicated Sluijter's study for American English, but with a change in focus. Instead of comparing full stressed vowels with primary stress to unstressed and reduced vowels, they compared unreduced vowels with primary stress to unreduced vowels with secondary stress. Their intention was to demonstrate that the patterns obtained by Sluijter and colleagues for American English were related to the vowel reduction differences between their target vowels. Campbell and Beckman hypothesized that the absence of vowel reduction would result in an absence of duration and spectral tilt differences related to stress. Their results confirmed their hypothesis: spectral balance did not differentiate levels of stress in the absence of a pitch accent, indicating that subjects did not use duration cues in a consistent fashion.

Thus, if Spanish patterns like Dutch, it will show stress differences based on duration and spectral tilt in deaccented contexts. If, on the other hand, it is true that in the absence of vowel reduction there are no differences between stress levels, as demonstrated for unreduced vowels in English, then Spanish, which has no vowel reduction, will not be able to maintain a stress contrast in contexts where there is no covariation between stress and accent. In the present study, in order to test these hypotheses, we will examine the phonetic cues of duration, vowel quality, intensity, and pitch movements in stressed and unstressed syllables as spoken within declarative sentences and parenthetic phrases in Spanish.

The article is organized as follows. Section 2 describes the methodology used for the production experiment. In Section 3, we present the main effects of pitch, duration, overall intensity and spectral tilt on the stress and accent dimension, as well as the results of the linear discriminant analyses. Finally, in Section 4, we discuss the relative strength of these four acoustic cues as correlates of stress and accent in Spanish and compare our results with the results for other languages.

2. *Methodology*

2.1 *Materials*

In order to examine the [+/−stress] contrast, we created a corpus of fifteen four-syllable verbs that end either in *-nimar*, like *desanimar* ("to discourage"), or in *-minar*, like *determinar* ("to determine"). As shown in Table 1, the target verbal forms used in the experiment had either a paroxytone stress in the present tense (i.e. *desa<u>ni</u>mo* "I discourage", *deter<u>mi</u>no* "I determine/calculate"), or an oxytone stress in the past tense (i.e. *desani<u>mó</u>* "(s)he discouraged" and *determi<u>nó</u>* "(s)he determined/calculated"). In this way, we were able to contrast syllables that have the same segmental content and that differ only in degree of prominence, for example, stressed [no] in *determi<u>nó</u>* vs unstressed [no] in *deter<u>mi</u>no,* and stressed [mi] in *deter<u>mi</u>no* vs unstressed [mi] in *determi<u>nó</u>*. [N.B. throughout the article, stressed syllables are underlined.]

Spanish verbs	Present, 1st person sing. (paroxytones)	Past, 3rd person sing. (oxytones)
abominar "to abominate"	abomino	abominó
determinar "to determine"	determino	determinó
denominar "to name"	denomino	denominó
desanimar "to discourage"	desanimo	desanimó
descaminar "to mislead"	descamino	descaminó
discriminar "to discriminate"	discrimino	discriminó
diseminar "to spread"	disemino	diseminó
eliminar "to eliminate"	elimino	eliminó
encaminar "to guide"	encamino	encaminó
examinar "to examine"	examino	examinó
exterminar "to exterminate"	extermino	exterminó
iluminar "to light"	ilumino	iluminó
incriminar "to incriminate"	incrimino	incriminó
predominar "to predominate"	predomino	predominó
recriminar "to recriminate"	recrimino	recriminó

Table 1: *Target verbs used in the experiment.*

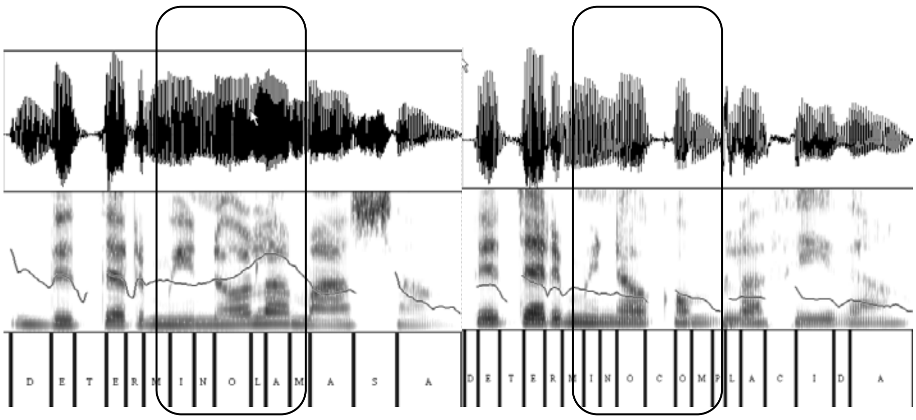

Figure 1: *Waveform, spectrogram with F0 track, and segmentation tier of the declarative utterance* Determinó la masa *"She determined the mass" (left) and of the quotation mark* Determinó complacida *"She determined in a satisfied way" (right).*

In order to control for stress and accent covariation, each of the fifteen four-syllable verbs was embedded in a segmentally identical utterance fragment that was spoken with either a declarative intonation or the flat intonation of parenthetic sentences. In declarative sentences, stressed syllables also bear a pitch accent while unstressed syllables remain deaccented. In contrast, in parenthetical intonation, F0 is flat across the utterance and shows no pitch accents (Figure 1).

Thus, for each verb, we obtain the four-sentence set shown in Table 2: one declarative sentence with the verb in the present tense in (a), one declarative sentence with the verb in the past tense in (b), one parenthetic sentence with the verb in the present in (c), and one parenthetic sentence with the verb in the past in (d). The data results in a total of three hundred syllabic tokens: two syllabic positions (final and penultimate) x two utterance types (declarative and parenthetic) x fifteen verbs x five subjects.

	Declarative sentences	Parenthetic sentences
[+stress] paroxytone verbs	[+accent] (a) *Deter**mi**no la masa.*	[−accent] (c) *—La masa del átomo es medible— deter**mi**no complacida.*
[−stress] oxytone verbs	[−accent] (b) *Deter**mi**nó la masa.*	[−accent] (d) *—La masa del átomo es medible— deter**mi**nó complacida.*

Table 2: *Target syllable* mi *(in bold) in four sentences. Underlining indicates stressed syllables.*

2.2 *Procedure*

Thirty cards were prepared which each showed a verb in infinitival form, a context, and two questions with their corresponding answers. The subjects were told that they would hear a question and should then read the appropriate answer with the corresponding intonation, i.e. either with a declarative intonation and or the flat intonation of a parenthetic sentence. After shuffling the thirty cards, the experimenter (either the first or the second author) chose the card on top of the pile and then read aloud the verb, context, and first question to the subject. The subject said the answer with the appropriate intonation. Then the second question was read and the subject read out the second answer accordingly. If the experimenter thought that the subject's pronunciation or intonation of an utterance was unnatural, the speaker was asked to repeat the sentence. The process was repeated for each one of the thirty cards.

Speakers were recorded individually in a quiet room, using a Sennheiser MKH20P48U3 omnidirectional condenser microphone and a Pioneer PDR609 digital CD-recorder. Speech samples were digitized at 32000 Hz in 16-bit mono, and target utterances were double-checked to make sure that they had been produced with the intended prosody.

2.3 *Subjects*

Five native speakers of Barcelona Spanish, two male and three female, participated in the experiment. Their ages ranged from twenty-six to forty-two years old. All subjects had earned university degrees and spoke an educated variety of their Spanish dialect. They reported that they normally spoke this language with their parents and siblings, and had learnt Catalan later as a second language in school. No subject reported having speech or hearing problems.

2.4 *Data Analysis and Measurements*

The following measurements were made with Praat (Boersma & Weekink 2005; Wood 2005) on each of the three hundred syllabic tokens.

2.4.1 *Fundamental Frequency.* We took the general view that pitch movements are the correlate of accent. In order to test this assumption, we measured the pitch range of the target pitch accent (in the accented case) or target syllable (in the unaccented case). The valleys and peaks of the prenuclear pitch accents (see L and H marks in Figure 2) were then marked. In the cases where the pitch was completely flat, such as in the parenthetic sentences, marks were placed at the beginning and at the end of the syllable. A Praat script extracted the F0 value in Hz at the marked points and calculated the pitch range by subtracting the F0 values at L from the F0 values at H for each of the three hundred tokens. Pitch range was given in absolute values.

2.4.2 *Duration.* Each segment of the verb endings *-mino* and *-nimo* was marked according to the F2 transitions displayed in the spectrograms. Vowels contained the transitions (see marks for 'm', 'i', 'n', 'o' in tier 1 of Figure 2). A script calculated the duration of each segment and each syllable in milliseconds.

2.4.3 *Vowel quality.* Formant measurements were based on the stable part of the vowel as marked in tier 3 in Figure 2. F1, F2 and F3 were calculated as frequency averages in Barks. Vowel quality for each token was computed as the difference between F2 and F1.

2.4.4 *Intensity*. Following Sluijter and van Heuven (1996b), we estimated intensity in terms of both overall intensity and spectral tilt (or spectral balance). Speakers were first normalized for overall differences in intensity. By using an algorithm included in the sound editing software 'Cool Edit', the loudest part of the waveform was set to a specified amplitude, −10dBFS in our study, thereby raising or lowering all other parts of the same waveform by the same amount. In this way, we ensured that all files and all speakers had a consistent volume.

Overall intensity was estimated using the command 'Get intensity' from Praat over the stable part of each vowel (tier 3 in Figure 2), after having levelled each sentence for loudness.

To obtain the measures of spectral tilt for vowel [o], we extracted the amplitudes of two frequency bands as segmented in tier 3: band 1 ranged from 0 to 400 Hz and band 2 from 400 Hz to 4000 Hz. Band 1 contained F0 while band 2 contained the vowel formants. The same procedure could not be performed on vowel [i], because F1 frequency was too low to be separated from F0. The spectral tilt for vowel [o] was computed as the ratio of band 2 to band 1. Thus a score closer to 1 indicates that the intensity from the lower frequencies is similar to that in the highest frequencies, while a score closer to 0 shows that the intensity of that vowel is concentrated in the lower band.

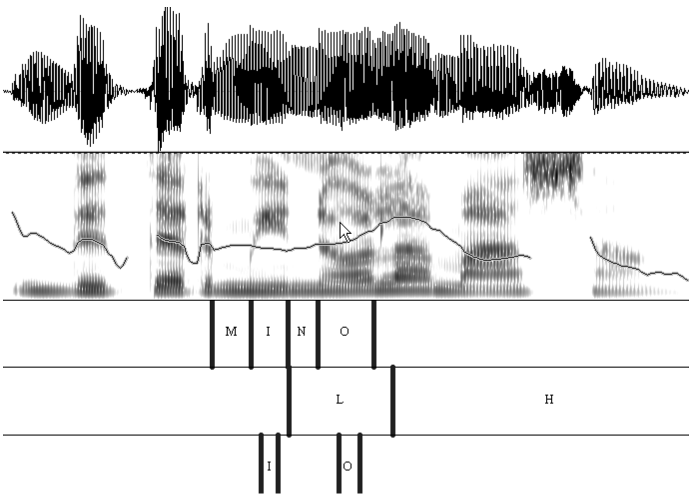

Figure 2: *Waveform, spectrogram, F0 trajectory, and segmentation tier of the declarative utterance* Determinó la masa *"(S)he calculated the mass"*.

2.4.5 *Statistical analysis.* We first performed a Repeated Measures ANOVA with the factors vowel ([i], [o]) and accent (+accent/−accent) on the measurements of pitch range in order to verify that accented and stressed vowels bore a pitch accent while unaccented and stressed vowels did not. As for duration, vowel quality, spectral tilt and overall intensity, we performed two statistical analyses. First, we ran a Repeated Measures ANOVA with stress (+/−stress) and intonation (declaratives/parenthetic sentences) as main factors on each vowel ([i] / [o]) for each set of measurements. Second, in order to investigate the contribution of each set of measurements in the prediction of stress we carried out a Linear Discriminant Analysis (LDA) with duration, vowel quality, spectral tilt, and overall intensity as the predictor variables and stress or accent as the criterion variables.

3. *Results*

3.1 *Pitch range differences*

One of the first things we wanted to check was whether accented syllables (in declarative sentences) were consistently produced with a rising pitch trajectory, in contrast with unaccented stressed syllables (in parenthetics), which were expected to be flat in pitch. The graph in Figure 3 shows mean values and standard error (in Hz) of the pitch range of the stressed syllables in paroxytones (in gray) and oxytones (in black) in accented and unaccented conditions for all five speakers. As is clear from the graph, subjects consistently used a pitch increase in declarative sentences (e.g. the [+accent] condition: mean 40.93 Hz, s.d. 30.22) and practically no increase or F0 variation in the parenthetic sentences (e.g. the [−accent] condition: mean −0.91, s.d. 4.11). A one-way ANOVA corroborated this difference as significant (F(1,298) = 526.222, p<0.0001). Thus, as expected, lexical stress was consistently cued by a pitch accent in declarative sentences, while it was not in parenthetic utterances.

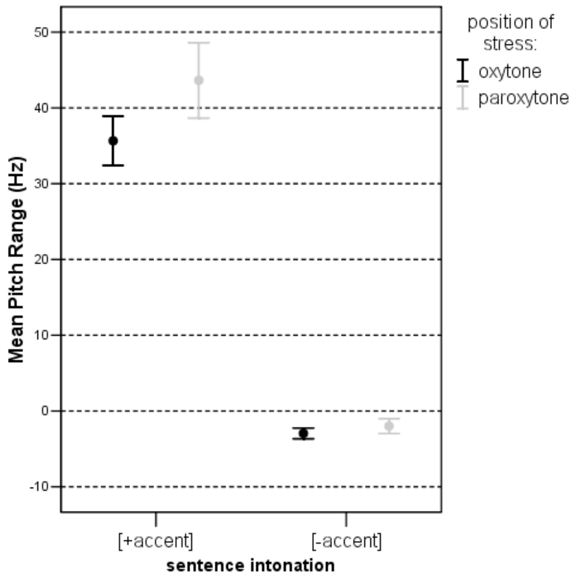

Figure 3: *Mean values and standard errors (in Hz) of the pitch range of stressed syllables.*

3.2 *Duration, vowel quality, overall intensity, and spectral tilt.*

3.2.1 *Duration.* The two graphs in Figure 4 plot the confidence intervals for the mean of the penultimate syllable *mi* (left panel) and word-final syllable *no* (right panel) in different stress (stressed/unstressed) and intonation (declarative/parenthetic) conditions for all five speakers. Three patterns stand out. First, stressed syllables (in grey) are systematically longer than unstressed syllables (in black), and most importantly, this difference is maintained across intonation contexts, meaning that stressed syllables are longer even in unaccented environments. Moreover, the magnitude of lengthening of the factor [stress] is greater in word-final syllables than in penultimate syllables (mean differences between stressed and unstressed syllables: 15 ms for word-final syllables vs 7 ms for penultimate syllables). Second, we find no consistent patterns with respect to the potential lengthening effects of accent: while word-final syllables (syllable 2) are longer in accented (declarative) contexts than in unaccented (parenthetic) contexts, this effect is not obvious for syllable 1 (mean differences between declarative and parenthetic sentences: 1 ms for penultimate syllables vs 6 ms for word-final syllables). Importantly, though, the magnitude of lengthening exerted by the presence of stress is higher than that produced by accent. Finally, the graphs in Figure 4 also show that final syllables are longer than penultimate syllables in all conditions (stressed, unstressed, accented, unaccented). This effect might be related to either the inherent duration of vowels or word position.

We ran a Repeated Measures ANOVA with the factors of stress (+stress/−stress) and intonation (declarative/parenthetic) on the duration of syllables 1 and 2. The main factor of stress was significant while the interaction 'stress x intonation' was non-significant, meaning that stressed syllables were longer than unstressed syllables in both conditions (stress: *[i]* $F(1,74) = 31.635$, $p<0.0001$; *[o]* $F(1,74) = 86.535$, $p<0.0001$; interaction: *[i]* $F(1,74) = 2.293$, $p = 0.134$; *[o]* $F(1,74) = 0.019$, $p = 0.891$). The main factor of intonation was significant only for vowel [o] (*[i]* $F(1,74) = 0.156$, $p = 0.694$; *[o]* $F(1,74) = 9.987$, $p = 0.002$).

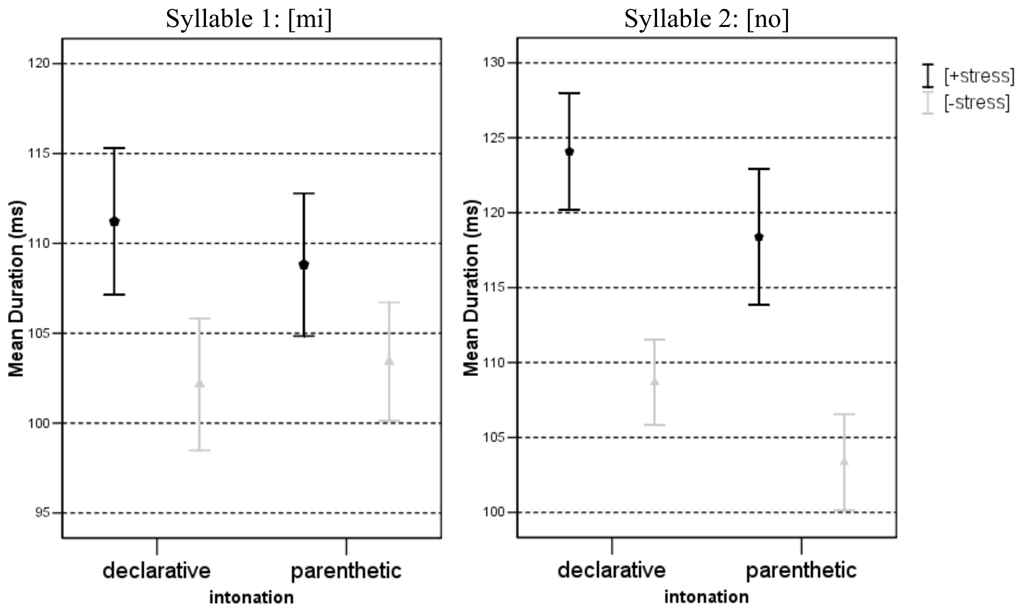

Figure 4: *Mean syllable duration of stressed and unstressed syllables in declarative and parenthetic utterances.*

Syllable	Intonation	Contrast	Significance
1: [mi]	Declarative	[+stress, −stress]	p<0.0001
	Parenthetic	[+stress, −stress]	p<0.01
2: [no]	Declarative	[+stress, −stress]	p<0.0001
	Parenthetic	[+stress, −stress]	P<0.0001

Table 3: *Results of paired-samples t-tests on duration of syllable 1 and syllable 2. Significance at 0.05 alpha level. Bonferroni adjustment for multiple comparisons.*

Table 3 shows the results of paired-samples t-tests comparing stressed with unstressed syllables within declarative and parenthetic sentences for each syllable. The results confirm that the duration differences between stressed and unstressed syllables remain significant within declarative and parenthetic sentences for both syllable 1 and syllable 2. Therefore, the differences in duration between stressed and unstressed syllables are significant regardless of the presence of an accent.

Further paired samples t-tests compared stressed accented syllables with stressed unaccented syllables, as well as unstressed and unaccented syllables from declarative and parenthetic sentences. As the results in Table 4 show, the differences between stressed accented and stressed unaccented syllables are only significant for syllable 2. Moreover, unstressed and unaccented syllables from declarative and parenthetic sentences are also significantly different only for syllable 2. These results confirm that the factor [+accent] does not yield a systematic additive effect on syllable duration and suggest that the factor involved in the lengthening of the last syllable is probably related to the inherent duration of vowels (i.e. the [i] in the penultimate syllable is shorter than the [o] in the last syllable), or within-word position (word-medial versus word-final position), rather than the property of being accented or not.

Syllable	Stress	Contrast	Significance
1: [mi]	[+stress]	[+accent, −accent]	p>0.05 n.s.
	[−stress]	[−accent, −accent]	p>0.05 n.s.
2: [no]	[+stress]	[+accent, −accent]	p<0.05
	[−stress]	[−accent, −accent]	p<0.01

Table 4: *Results of paired-samples t-tests on duration of syllable 1 and syllable 2. Significance at 0.05 alpha level. Bonferroni adjustment for multiple comparisons.*

We conclude on the basis of these results that duration is a strong acoustic correlate of the *stress* difference in Spanish, but not of the presence of an *accent*. In our data, the presence of an accent does not obligatorily trigger lengthening on the stressed syllable.

3.2.2 *Vowel quality*. Given that males tend to have lower formant values than female speakers, vowel quality was measured as the distance in Barks between F2 and F1 separately for female and male speakers in our data. The graphs in Figure 5 illustrate the mean confidence intervals for the mean F2-F1 difference in Barks for vowel [i] (female speakers in top left panel, male speakers in top right panel) and for vowel [o] (the panels at the bottom, females left panel, males right panel) in different stress (stressed/unstressed) and intonation conditions (declarative/parenthetic). The graphs reveal that the differences between stressed and unstressed vowels are less than 1 Bark for all the contexts. The direction of the reduction is consistent only for vowel [o]. Female and male speakers tend to increase the distance between F1 and F2 in unstressed [o], making it closer to a central vowel.

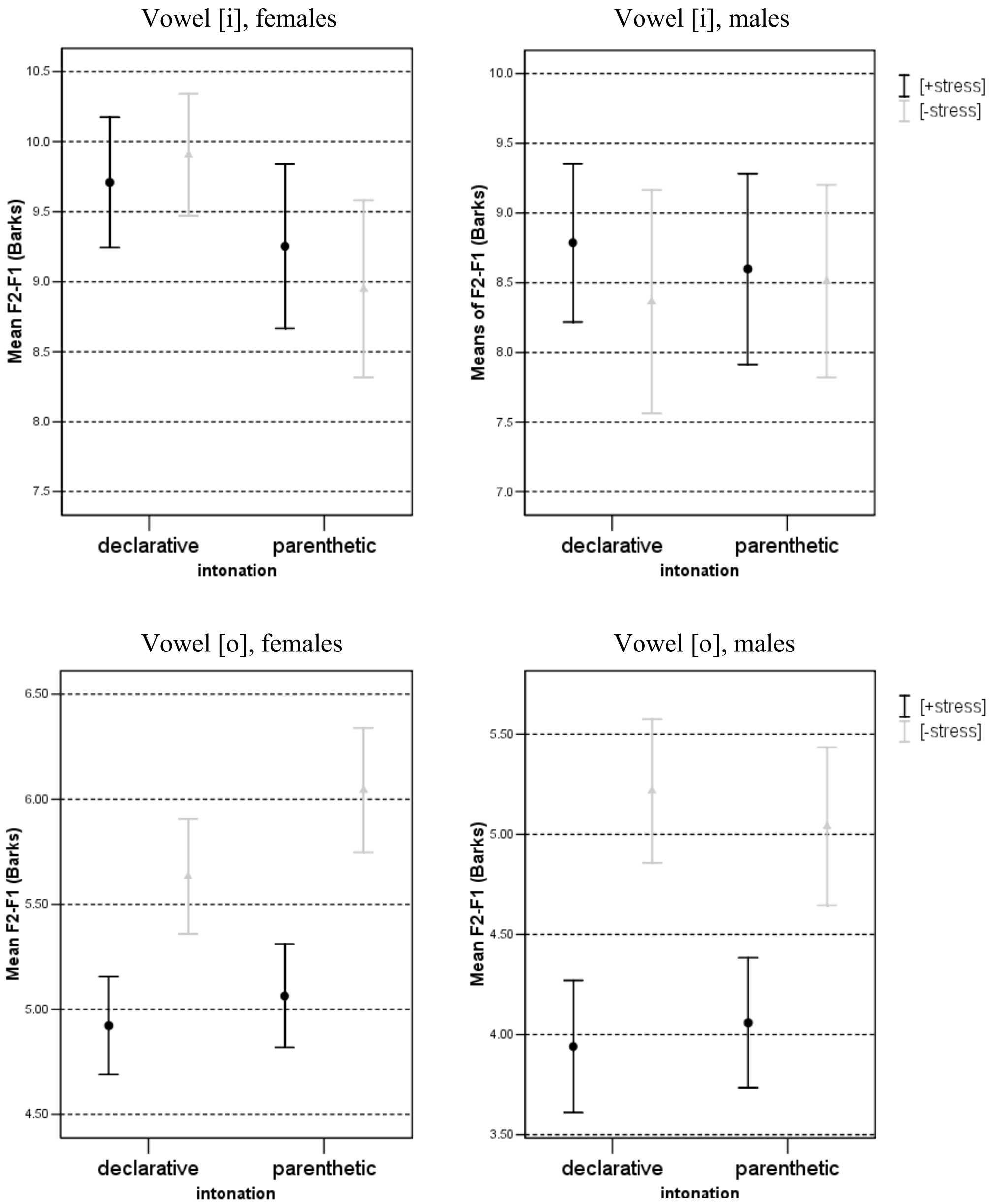

Figure 5: *Mean F2-F1 difference (in Barks) for vowel* [i] *and vowel* [o] *in different stress and intonation conditions.*

Repeated measures ANOVA with the factors of stress (+/−stress) and intonation (declarative/parenthetic sentences) for vowel [o] shows the main factor of stress to be significant for both female speakers ($F(1,44) = 25.098$, $p<0.0001$) and male speakers ($F(1,29) = 30.856$, $p<0.0001$), while intonation and the interaction 'stress x intonation' are non-significant. This indicates that only stress, and not accent, has an effect on vowel quality changes (intonation: females $F(1,44) = 0.288$, $p = 0.594$; males $F(1,29) = 0.083$, $p = 0.775$; interaction: females $F(1,44) = 0.146$, $p = 0.704$; males $F(1,29) = 0.300$, $p = 0.588$). Paired-samples t-tests confirm that the 1 Bark difference between stressed and unstressed [no] is maintained across intonation contexts by both female and male speakers.

Syllable [no]	Intonation	Contrast	Significance
Females	Declarative	[+stress, −stress]	p<0.0001
	Parenthetic	[+stress, −stress]	p<0.0001
Males	Declarative	[+stress, −stress]	p<0.0001
	Parenthetic	[+stress, −stress]	p<0.05

Table 5: *Results of paired-samples t-tests on vowel quality of syllable* [o]. *Significance at 0.05 alpha level. Bonferroni adjustment for multiple comparisons.*

In sum, we found a small but consistent effect of stress on the formant values of [o]: Unstressed [o] becomes slightly more centralized than stressed [o]. These results are in agreement with those of Quilis and Esgueva (1983), which showed a slight tendency for centralization in unstressed mid-vowels in Castilian Spanish, and with patterns of vowel reduction in unstressed syllables across Romance languages. In contrast, pitch accents did not have any significant effect on vowel quality changes.

3.2.3 *Overall intensity*. The two graphs in Figure 6 display the confidence intervals for the mean overall intensity (in dB) for vowel [i] (left panel) and vowel [o] (right panel) in different stress (stressed/unstressed) and intonation conditions (declarative/parenthetic) for all five speakers. In the first place, the graphs reveal that in contrast with duration and vowel quality, stressed and unstressed vowels differ in overall intensity only within declarative sentences. In parenthetic sentences, these differences tend to disappear. This means that, on the one hand, there is no consistent effect of stress on overall intensity measurements; on the other, it indicates a possible effect of accent. Moreover, note that the declarative 'stressed-unstressed' pattern in vowel [i] is reversed in vowel [o]. While stressed [i] has a higher overall intensity than unstressed [i], unexpectedly, for [o], it is the unstressed vowel that displays a higher overall intensity score.

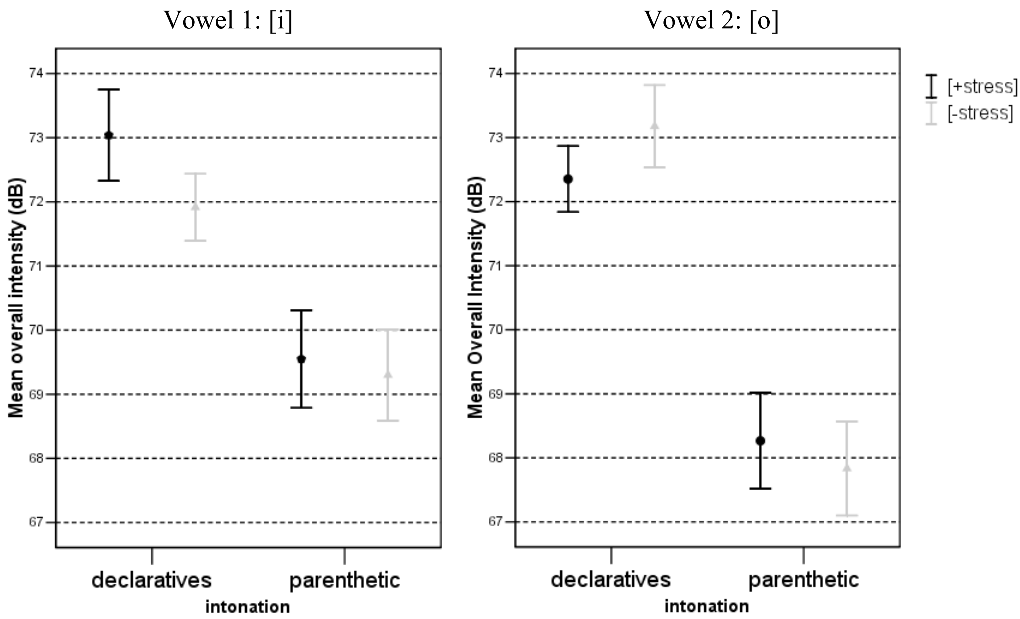

Figure 6: *Mean overall intensity and standard error (in dB) for vowel* [i] *(left panel) and vowel* [o] *(right panel) in different stress (stressed/unstressed) and intonation contexts (declarative/parenthetic) for all five speakers.*

Results from the Repeated Measures ANOVA show that the 'stress x intonation' interaction is significant for both syllables (vowel [i] $F(1, 74) = 27.140$, $p<0.0001$; vowel [o] $F(1, 74) = 20.559$, $p<0.0001$) indicating that patterns of overall intensity differ across declarative and parenthetic sentences. Paired-samples t-tests indicate that the difference in overall intensity is only active in declarative sentences. Since declarative sentences differ in regards to accent while parenthetic sentences do not, overall intensity relates to accent, not stress.

Vowels	Intonation	Contrast	Significance
1: [i]	Declarative	[+stress, −stress]	$p<0.0001$
	Parenthetic	[+stress, −stress]	$p>0.05$ n.s.
2: [o]	Declarative	[+stress, −stress]	$p<0.0001$
	Parenthetic	[+stress, −stress]	$p>0.05$ n.s.

Table 6: *Results of paired-samples t-tests on overall intensity of vowels 1 and 2. One-tailed, significance at 0.05 alpha level. Bonferroni adjustment for multiple comparisons.*

There is an asymmetry in the overall intensity values of declarative sentences: in vowel [i], unstressed syllables display lower overall intensities than their stressed counterparts, while in [o] this pattern reverses. We may be able to explain this if we consider the F0 trajectories of each vowel in detail and assume that there is a possible covariation between F0 trajectories and overall intensity. Greater intensity is generally found in accented syllables due to the larger amplitude of vocal fold vibration related to greater speaker effort (Sluijter & van Heuven 1996b:2472). All unstressed instances of [i]—i.e. [i] in parenthetic sentences (see Figure 1) and the unstressed [i] in declarative sentences—display a flat F0 trajectory. These vowels also show the lowest intensity values. By contrast, stressed accented [i]s bear a rising F0 trajectory, and, correspondingly, show the highest intensity values. As for [o], this vowel in parenthetic sentences displays a flat intonation contour and thus has lower intensity values than in declarative utterances, where it has a rising F0 trajectory. The stressed accented [o] in declarative sentences bears the pitch accent and therefore has both a rising F0 trajectory and high intensity values. Crucially, although unstressed [o]s in declarative sentences are phonologically unaccented, they bear the peak of the preceding pitch accent and display higher intensity values than unstressed [o]s at the beginning of the F0 rising trajectory. In fact, the results in Table 7 below demonstrate that all subjects showed a significant positive correlation between overall intensity and F0 height for each vowel, the correlation coefficients being especially strong for subjects 2, 3, and 4. That is, the higher the pitch of the vowel, the higher the overall intensity levels obtained.

Therefore, it is hypothesized that the increased overall intensity patterns found in the Spanish data are due to the interdependence between F0 levels and overall intensity. As Sluijter and van Heuven (1996b:2482) claim, the greater intensity typically found in accented syllables is caused by the larger amplitude of the pulses in vocal fold vibration.

	vowel [i]	vowel [o]
Subject 1	0.507	0.552
Subject 2	0.713	0.641
Subject 3	0.751	0.750
Subject 4	0.783	0.721
Subject 5	0.586	0.635

Table 7: *Correlation coefficients between overall intensity and F0 height for the two vowels* [i] *and* [o] *for five subjects. One-tailed, all cells were significant at 0.01 level.*

3.2.4 *Spectral tilt*. In our data, spectral tilt was calculated as the ratio of the intensity in the higher frequencies to the intensity of the lower frequencies in vowel [o], as spectral tilt could not be measured for vowel [i] (see Section 2.3.3). Thus, when frequencies from the higher and from the lower part of the spectrum have similar intensities, the ratio approaches 1 and the tilt in the spectrum decreases. Figure 7 shows the mean spectral tilt ratios (and standard error values) for vowel [o] in different stress (stressed/unstressed) and intonation conditions (declarative/parenthetic) for all five speakers. First, the spectral tilt ratios of stressed [o]s (in grey) are closer to 1 and show a flatter tilt than unstressed [o]s (in black). Like for duration and vowel quality, this difference is maintained across intonation contexts, revealing a potential effect of stress on spectral tilt: stressed syllables tend to increase the intensity of the higher frequencies, and consequently have a ‘flatter’ spectral tilt than their unstressed counterparts. Second, the spectral tilt of [o] in declarative sentences is closer to 1, and therefore the tilt decreases, in contrast to parenthetic sentences. This reveals a potential effect of the presence of an accent.

Results on the Repeated Measures ANOVA show that the interaction ‘stress x intonation’ is non-significant for spectral tilt measurements ($F(1,74) = 1.797$, $p = 0.185$ for vowel [o]), indicating that the effect of stress on these independent variables is the same regardless of the presence of an accent. Moreover, paired-samples t-tests confirm that spectral tilt is a reliable acoustic correlate of stress across intonation contexts.

Paired T-tests show that there is a significant difference between accented and unaccented syllables, and between unaccented syllables from declarative and parenthetic sentences. These results indicate that there is a difference between sentence type: declarative sentences display greater intensity levels in the higher regions of the spectrum than parenthetic sentences.

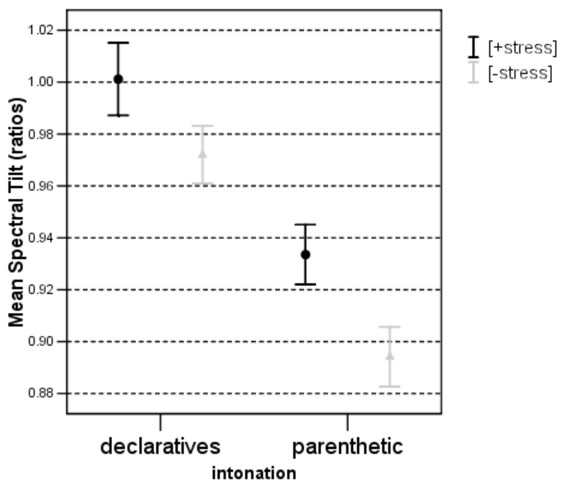

Figure 7: *Mean spectral tilt ratios (and standard error values) for vowel* [o] *in different stress (stressed/unstressed) and intonation (declarative/parenthetic) contexts for all five speakers.*

Vowel	Intonation	Contrast	Significance
2: [o]	Declarative	[+stress, −stress]	p<0.05
	Parenthetic	[+stress, −stress]	p<0.0001

Table 8: *Results of paired-samples t-tests on spectral tilt of vowel* [o]. *One-tailed, significance at 0.05 alpha level. Bonferroni adjustment for multiple comparisons.*

These results therefore suggest that spectral balance is a more robust and systematic cue to stress than overall intensity, and are in keeping with previous results on other stress-accent languages (Sluijter & van Heuven 1996a, 1996b). We thus suggest that Navarro Tomás's hypothesis that Spanish stress is strongly cued by intensity (the so-called *acento de intensidad*) can be interpreted as essentially correct if one understands that the acoustic correlate of increased perception of loudness is greater intensity levels in the higher parts of the spectrum. Thus the perception that a stressed syllable is more prominent probably derives from its increased intensity levels in the high regions, not the low regions, of the spectrum. This difference is maintained in unaccented contexts and enhanced in accented syllables.

Vowel	Stress	Contrast	significance
2: [o]	[+stress]	[+accent, −accent]	p<0.0001
	[−stress]	[−accent, −accent]	p<0.0001

Table 9: *Results of paired-samples t-tests on spectral tilt of vowel* [o]. *One-tailed, significance at 0.05 alpha level. Bonferroni adjustment for multiple comparisons.*

3.2.5 *Linear Discriminant Analyses.* Following Sluijter and van Heuven (1996b), the contribution of each acoustic correlate was examined by Linear Discriminant Analyses (LDA). Two LDA with the grouping variable of stress (+stress/−stress) were performed on measurements of duration, vowel quality, spectral tilt and overall intensity for vowel [o]. Since spectral tilt could not be measured for vowel [i], LDA was not run for this vowel (see Section 2.3.3). First, all measurements were entered together in order to assess how well stress could be predicted. The obtained discriminant functions correctly classify as [+stress] or as [−stress] 71.3% of the vowel [o] tokens in declarative sentences and 70.7% in parenthetic sentences. Thus, stress could be predicted with a reasonable level of accuracy from the measurements of duration, spectral tilt, vowel quality and overall intensity.

Secondly, measurements of duration, spectral tilt, vowel quality and overall intensity were entered separately into the discriminant function so as to determine the contribution of each one of these variables in the prediction of stress. As Figure 8 shows, duration correctly classified 70% of stressed syllables in declarative sentences and 66.7% in parenthetic sentences. Vowel quality classifications achieved scores of 60.7% in declaratives sentences and 57.3% in parenthetic sentences. Spectral tilt scored 51.3% in declarative sentences but increased to 61% in parenthetic sentences. In contrast with the preceding measurements, classification scores for overall intensity were always below chance (46% in declarative sentences and 50% in parenthetic sentences).

These results indicate that *duration* is the main cue to stress in Spanish. Duration measurements showed that stressed syllables had longer durations than their unstressed counterparts, and the LDA results indicate that these differences in duration are sufficient to distinguish stressed from unstressed syllables with a high level of accuracy. Moreover, since the scores for declarative sentences do not vary substantially from those obtained for parenthetic sentences (3.3% difference), they show that the successful classification of stressed syllables based on duration differences takes place regardless of the presence or absence of a pitch accent. Thus, duration does effectively cue the stress contrast independently of pitch accents.

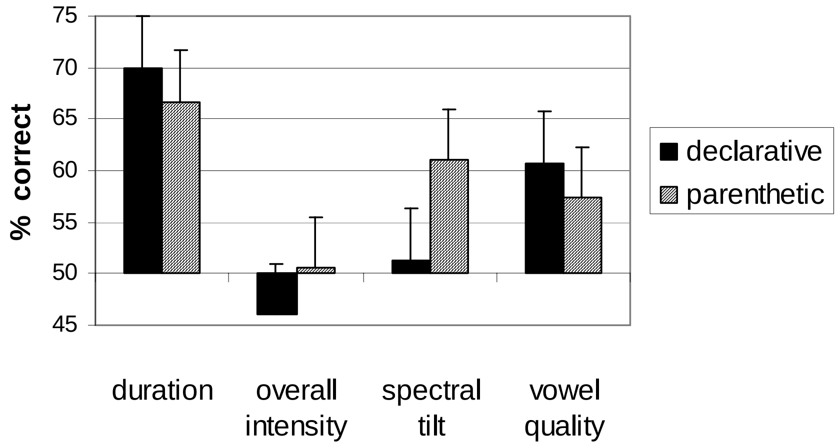

Figure 8: *Percentage of vowels correctly predicted as stressed or unstressed for each phonetic cue.*

A similar pattern is found for *vowel quality*: the distance between F1 and F2 increases slightly in unstressed [o] in both declarative and parenthetic sentences, indicating a slight tendency towards vowel centralization in unstressed vowels, which in turn leads to correct classification scores of stressed syllables in both sentence types. This tendency to vowel centralization, however, is based on a difference of less than 1 Bark, which may call into question the perceptual relevance of this cue.

Spectral tilt also contributes to the prediction of stress, but only in parenthetic sentences. This may indicate a compensatory relation between duration (and possibly vowel quality) and spectral tilt. Since duration cues have less predictive power in parenthetic sentences, spectral tilt becomes a better predictor of stress in this context. The only cue that does not contribute to the prediction of stress in any context is *overall intensity*.

In summary, LDA of the Spanish data show that duration is the most effective correlate of stress in both parenthetic and declarative sentences. After that, vowel quality makes a significant contribution in predicting vowels as [+stressed] or as [−stressed], followed by spectral tilt. Overall intensity, however, does not contribute to this classification.

4. *Discussion and conclusions*

In this article, we were concerned with the acoustic correlates that characterize stress and accent in Spanish. We analyzed four acoustic correlates of stress (syllable duration, vowel quality, overall intensity, and spectral balance) in four conditions, namely, stressed and unstressed syllables in both accented and unaccented environments. This allowed us to examine the relative strength of these correlates with relation to stress and see how the stress contrast is maintained in the presence or absence of a pitch accent.

The duration measurements revealed that stressed syllables are longer than unstressed syllables regardless of the presence of an accent, demonstrating that syllable duration is a strong acoustic correlate of the stress difference in Spanish. Moreover, LDA results for the Spanish data singled out the effectiveness of duration as the most robust acoustic separator between stressed and unstressed conditions. This is basically in accordance with the main results for Dutch, where duration is the most effective correlate of stress (Sluijter & van Heuven 1996a, 1996b:2475). Furthermore, in contrast with previous studies, our results show that the presence of a pitch accent does not consistently trigger additive effects on the duration cues. That is, in our data, the presence of an accent does not obligatorily trigger lengthening on the stressed syllable. Even though previous studies on other stress-accent languages have found additive effects of accent (Sluijter & van Heuven 1996a, 1996b:2475, for English and Dutch respectively), Beckman and Edwards (1994:20-25) found that this pattern varied across speakers and speech rates: while one of the speakers showed a consistent durational effect of accent, this was not the case for the other speaker. Consequently, we claim, along with Beckman & Edwards (1994), that while duration is a crucial acoustic cue to mark a lower level prominence contrast (stressed vs unstressed), it is a secondary (and thus optional) acoustic marker of a higher-level prominence contrast (accented vs unaccented).

The formant measurements for the Spanish data revealed significant effects of stress on the formant values of [o], indicating a slight tendency towards centralization in unstressed positions. Moreover, LDA results confirm the significant contribution of the vowel quality variable to stress prediction (albeit less strong than duration). These results contrast with those for Dutch, where the effects of stress on vowel quality were only partially significant and vowel quality was found to be a poor predictor of stress in LDA analyses (Sluijter & van Heuven 1996b). On the one hand, this difference could be related to the fact that Sluijter used both vowels in the LDA prediction ([a] and [ɔ]) and only [a] reduced into [ɑː] in unstressed syllables. On the other hand, vowel reduction in Dutch takes place mainly in derivational suffixes (van Heuven 2001), which were not examined in Sluijter's test materials. In our data, we included only [o], which underwent vowel quality changes, not [i], which did not. If we had included both vowels in the LDA analysis, vowel quality might not have been such a good predictor of stress, and our results might have been closer to those of Sluijter and van Heuven.

On the other hand, the presence of an accent does not affect formant frequency values, and therefore accented syllables have similar vowel qualities to unaccented syllables. Thus, both syllable duration and vowel quality cues can be interpreted as 'primary' cues in the stress dimension and 'secondary' cues in the accent dimension. This is probably due to the fact that in our data vowel lengthening was related to stress, not to accent, and as Lindblom showed (Lindblom 1963; Moon & Lindblom 1994), there is a linear relationship between duration and formant displacement: shorter vowels undergo more formant displacement (towards centralization) than longer vowels. This linear relationship between duration and formant displacement is biomechanically motivated and provides evidence for a vowel undershoot model. In shorter vowels, articulators have less time to attain their target, and as a result, vowels become reduced, thus showing more formant displacement towards a reduced vowel. Thus, unstressed vowels in Spanish become slightly more centralized than stressed vowels because they are also shorter than their stressed counterparts. However, the magnitude of this centralization is very small, probably because Spanish does not have phonological vowel reduction. It would be interesting to compare Spanish to a language with phonological vowel reduction in order to examine how these two variables cue the stress contrast in each language.

We turn now to a discussion of intensity patterns. Crucially, the data presented in this article replicates Sluijter and collaborators' (1996a, 1996b) finding that the intensity differences between stressed and unstressed vowels are mainly located in the higher regions of the spectrum and, as Campbell and Beckman (1997) showed, these differences are enhanced in accented contexts. It is clear that overall intensity cannot be regarded as a reliable acoustic correlate of stress in Spanish, as our ANOVA and LDA results demonstrate. By contrast, spectral balance differences (i.e. intensity levels at higher regions of the spectrum), appear to be a consistent cue for stress. Thus, we contend that the classic claim in the Spanish phonetics literature made by Contreras (1963) and Quilis (1971, 1981) that intensity plays almost no role as a cue to stress is not accurate. On the other hand, we take Navarro Tomás' view that intensity plays an essential part in the production of stress (the so-called *acento de intensidad*) as essentially correct. Thus, a Spanish stressed syllable is probably perceived as more prominent due to an increase in the intensity levels in the higher, not lower, regions of the spectrum.

The present findings also have implications for previous debates on acoustic correlates of Spanish stress, as this is one of the first experiments comparing the acoustic correlates of Spanish stress in accented and unaccented words. Previous studies on the acoustic characterization of Spanish stress had only studied words containing a pitch accent (see Navarro Tomás 1914, 1964; Contreras 1963, 1964; Quilis 1971; Gili Gaya 1975; Solé 1984; Canellada & Kuhlman-Madsen 1987; Llisterri et al. 2003, among many others). In this new context, the traditional goal of searching for the *main cue* to stress in Spanish makes no sense, as phonetic cues are not used the same way in accented and unaccented contexts. In accented contexts, it is clear that pitch is a strong phonetic cue of stress, as claimed by Contreras (1964) or Quilis (1971, 1981) and authors of perception studies like Solé (1984), Enríquez, Casado & Santos (1989), and Llisterri et al. (2003); in this context, duration and intensity cues also accompany the pitch difference. Yet in unaccented contexts, where pitch is flat and cannot be an indicator of the stress difference, the results of this study reveal that duration, intensity, and even vowel quality are good indicators of the stress difference. In this sense, the traditional claim by Navarro Tomás that the strongest cue to stress in Spanish is a local increase in loudness or intensity is only partially true, as clearly duration is also a very strong indicator of the presence of stress. Thus in the absence of an accent, cues like duration and spectral tilt are crucial in the production of Spanish stress.

We conclude on the basis of these results that syllable duration, vowel quality, and spectral tilt (intensity at high frequencies of the spectrum) are all reliable acoustic correlates of the stress difference in Spanish. Accentual differences are acoustically marked by intensity cues, but our findings cast doubt on the notion that these might be a by-product of higher F0 levels which covary with higher intensity levels. Thus, our results reveal that American English, Dutch and Spanish do differ fundamentally in the use of vowel reduction and consonant reduction (flapping, aspiration) to mark stressed positions, but do not differ greatly in the way they use the other acoustic correlates (duration and intensity) to signal the presence of stress and accent. Stress is cued by duration, intensity, and vowel quality in the absence of an accent, confirming the relative independence of metrical and pitch properties. Finally, an appropriate follow-up of this research would be to examine the relevance and interaction of these factors in the actual perception of the stress contrast in Spanish.

References

Beckman, Mary E. 1986. *Stress and Non-stress Accent*. Dordrecht: Foris.

---------- & Jan Edwards. 1994. "Articulatory evidence for differentiating stress categories". *Phonological Structure and Phonetic Form. Papers in Laboratory Phonology III* ed. by Patricia A. Keating. 7-33. Cambridge: Cambridge University Press.

Boersma, Paul & David Weenink. 2005. *Praat: Doing Phonetics by Computer* (Version 4.3.01). [Computer program]. Retrieved from http://www.praat.org/

Bolinger, Dwight L. 1958. "A theory of pitch accent in English". *Word* 14. 109-149.

Canellada, María Josefa & John Kuhlman-Madsen. 1987. *Pronunciación del español. Lengua hablada y literaria*. Madrid: Editorial Castalia.

Campbell, Nick and Mary Beckman. 1997. "Stress, prominence and spectral tilt". *Intonation: Theory, Models and Applications. Proceedings of an ESCA Workshop* ed. by Antonis Botinis, Georgios Kouroupetroglou & George Carayiannis. 67-70. Athens: ESCA & University of Athens Department of Informatics.

Contreras, Heles. 1963. "Sobre el acento en español". *Boletín del Instituto de Filología de la Universidad de Chile* 15. 223-237.

----------. 1964. "¿Tiene el español un acento de intensidad?" *Boletín del Instituto de Filología de la Universidad de Chile* 16. 237-239.

Enríquez, Emilia V., Celia Casado & Andrés Santos. 1989. "La percepción del acento en español". *Lingüística Española Actual* 11. 241-269.

Gili Gaya, Samuel. 1975. *Elementos de fonética general*. Madrid: Gredos.

Heuven, Vincent J. van. 2001. Boven de klanken / Beyond the segments. Inaugural address in acceptance of the chair in Phonetics at Leiden University, The Netherlands. Retrieved from: http://www.let.leidenuniv.nl/ulcl/faculty/vheuven/oratie-eng.htm.

Lindblom, Björn. 1963. "Spectrographic study of vowel reduction". *Journal of the Acoustical Society of America* 35. 1773-1781.

Llisterri, Joaquim, María Machuca, Carme de la Mota, Montserrat Riera & Antonio Ríos. 2003. "The perception of lexical stress in Spanish". *Proceedings of the 15th International Conference of Phonetic Sciences* ed. by Maria-Josep Solé, Daniel Recasens & Joaquín Romero. 2023-2026. Barcelona: Causal Productions.

Navarro Tomás, Tomás. 1926 [1914]. *Manual de pronunciación española.* 2nd ed. Madrid: Hernando.

----------. 1964. "La medida de la intensidad". *Boletín del Instituto de Filología de la Universidad de Chile* 16. 231-235.

Moon, Seung-Jae & Björn Lindblom. 1994. "Interaction between duration, context, and speaking style in English stressed vowels". *Journal of the Acoustical Society of America* 96:1. 40-55.

Ortega-Llebaria, Marta. 2006. "Phonetic cues to stress and accent in Spanish". *Selected Proceedings of the 2nd Conference of Laboratory Approaches to Spanish Phonology* ed. by Manuel Díaz-Campos. 104-118. Somerville, Mass.: Cascadilla Press.

Pierrehumbert, Janet B. 1980. The phonetics and phonology of English intonation. PhD. diss., MIT.

Quilis, Antonio. 1971. "Caracterización Fonética del Acento Español". *Travaux de Linguistique et de Littérature* 9. 53-72.

----------. 1981. *Fonética acústica de la lengua española*. Madrid: Biblioteca Románica Hispánica, Gredos.

---------- & Manuel Esgueva. 1983. "Realización de los fonemas vocálicos españoles en posición fonética normal". *Estudios de Fonética* I ed. by Manuel Esgueva & Margarita Cantarero. 159-252. Madrid: Consejo Superior de Investigaciones Científicas.

Sluijter, Agaath M.C. & Vincent van Heuven. 1996a. “Acoustic correlates of linguistic stress and accent in Dutch and American English”. *Proceedings of the 4th International Conference on Spoken Language Processing* ed. by H. Timothy Bunnell & William Idsardi. 630-633. New Castle, Del.: Citation Delaware.

Sluijter, Agaath M.C. & Vincent van Heuven. 1996b. “Spectral balance as an acoustic correlate of linguistic stress”. *Journal of the Acoustical Society of America* 100:4. 2471-2485.

Sluijter, Agaath M.C., Vincent van Heuven & Jos J.A. Pacilly. 1997. “Spectral balance as a cue in the perception of linguistic stress”. *Journal of the Acoustical Society of America* 101:1. 503-513.

Solé, Maria-Josep. 1984. “Experimentos sobre la percepción del acento”. *Estudios de Fonética Experimental* I. 134-243. Universidad de Barcelona.

Wood, Sidney. 2005. *Praat for Beginners*. [Manual]. Retrieved from http://www.ling.lu.se/persons/Sidney/praate/