

# Acoustic Correlates of Stress in Central Catalan and Castilian Spanish

Language and Speech  
54(1) 73–97

© The Author(s) 2010

Reprints and permissions:

sagepub.co.uk/journalsPermissions.nav

DOI: 10.1177/0023830910388014

las.sagepub.com



**Marta Ortega-Llebaria**

University of Pittsburgh

**Pilar Prieto**

Institució Catalana de la Recerca i Estudis Avançats and Universitat Pompeu Fabra

## Abstract

The general literature on the phonetic correlates of stress agrees that duration, and in stress accent languages,  $F_0$  are consistent correlates of stress. However, the role of amplitude changes in the speech signal is more controversial. In particular, the conflicting results of spectral tilt as a correlate of stress have been attributed to the effects of vowel reduction. We examined the stress correlates of duration, overall intensity and spectral tilt in Catalan and Spanish in both accented and unaccented contexts while controlling for formant frequency differences between morphologically corresponding vowels in stressed and unstressed environments by comparing vowels that maintain the same quality across stress contexts with those that do not. Duration was a consistent stress correlate in all vowels in both languages, regardless of their formant frequency differences across stress contexts and of the absence of pitch accents. In fact, stress-related formant frequency differences between corresponding vowels amplify the duration cues to the stress contrast. On the other hand, the use speakers made of intensity was not as pervasive as that of duration. Specifically, changes in spectral tilt were significant only in Catalan and in those vowels that alternate a more open and peripheral realization in stressed syllables with a mid-central realization in unstressed syllables, indicating that spectral tilt is related to the formant frequency differences linked to the centralization processes rather than to the stress contrast.

## Keywords

accent, Catalan, prosody, Spanish, stress

## Introduction

Although object (the noun) and object (the verb) have the same segmental structure, their acoustic realizations are different. Fry (1955, 1958, 1964) showed in his pioneering studies that the stressed

---

### Corresponding author:

Marta Ortega-Llebaria, Department of Linguistics, University of Pittsburgh, Pittsburgh, PA 15260, USA

Email: mllebaria@yahoo.com

syllables ‘ob’ in the noun and ‘ject’ in the verb are realized with longer durations, louder intensities and more peripheral vowel qualities than their unstressed counterparts. Moreover, since stressed syllables in English attract pitch accents, that is, the F0 movements that configure the intonation of a sentence, English listeners also use the timing between the syllable and the pitch accent to convey stress prominence. Thus, the stress contrast in words like ‘object’ (the noun) and ‘object’ (the verb) is expressed by the suprasegmental parameters of F0, duration and signal amplitude, which can vary independently while the segmental identity of a syllable remains unaltered.

Research over the last fifty years has not overturned in fundamental ways Fry’s claims of F0 and duration as stress correlates. Early work on stress accent languages has examined stress correlates in sentences spoken with declarative intonation, where stressed syllables align with pitch accents and unstressed syllables do not, causing covariation between stress and accent (for English, see Fry, 1955, 1958; Lehiste and Peterson, 1959; Lieberman, 1960; Sweet, 1906; for Spanish, see Canellada and Kuhlman-Madsen, 1987; Contreras, 1963, 1964; Enríquez, Casado, and Santos, 1989; Gili Gaya, 1975; Llisterri, Machuca, de la Mota, Riera, and Ríos, 2003; Navarro-Tomás, 1914, 1964; Quilis, 1971, 1981; and Solé, 1984, among others; for Catalan, see Barnils, 1933; and Recasens, 1986). In these studies, F0 excursions were found to correlate with stress. Not all stressed syllables, however, 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. Therefore, in unaccented sentences, like the post-focal contexts examined in Huss (1978) and Nakatani and Aston’s (1978) experiments, F0 loses its effect as a stress correlate.

Beckman and Edwards (1994) expanded Huss’ (1978) and Nakatani and Aston’s (1978) results by examining words like *papa*, which contain an open stressed vowel and a central unstressed schwa, in focal and post-focal contexts, (a) and (b) respectively. They examined duration and articulatory correlates of stress in three target syllables. First, as in Huss (1978), they compared open vowels that differed in bearing a pitch accent (or not), like stressed *pa* in the first and second sentence. Second, they compared unaccented vowels that differed in vowel quality, i.e., open vowel versus central schwa, like the first and second *pa* in the second sentence.

- (1a) [Was her mama a problem about the wedding?]  
Her PAPA posed a problem
- (1b) [Did his dad pose a problem as far as their getting married?]  
HER papa posed a problem

Beckman and Edwards found that speakers consistently produced open stressed vowels with longer durations and larger and faster movements than their unstressed and centralized counterparts, while only some speakers used the same correlates to distinguish between stressed open vowels that contrasted in accent. Thus, duration and articulatory movements were reliable correlates of the stress contrast between the open and central vowels of a word, while the stress contrast between open vowels of different words was conveyed mainly by their association with distinct pitch patterns, showing that duration is a consistent correlate of stress at the word level and pitch accents are stress correlates at the sentence level.

Studies in other languages examined word stress in unaccented contexts similar to that in 1b (Sluijter and van Heuven, 1996a; Sluijter, van Heuven, and Pacilly, 1997, for Dutch; Manolescu, Olson, and Ortega-Llebaria, 2009, for Romanian; Dogil and Williams, 1999, for German; Kastriani, 2003, for Greek; Ortega-Llebaria, 2006, for Spanish). They also found that stressed syllables were consistently longer than their unstressed counterparts in the absence of pitch accents, revealing that duration was a cross-linguistic correlate of stress.

In contrast with F0 and duration, there is no general consensus on the role of signal amplitude as a cross-linguistic correlate of word stress (see Cutler, 2005, for a review). On the one hand, there is the common finding that amplitude manipulations have a weak effect on stress perception. For example, Turk and Sawusch (1996) showed that although duration and intensity were processed as a unit, irrelevant duration variations had a larger effect on the perception of loudness than irrelevant intensity variations in the perception of length, demonstrating that in English, either duration or duration and intensity can be cues to stress but not intensity alone.

On the other hand, several measures of intensity have been found to be consistent stress correlates. Beckman (1986) proposed that a measure of total amplitude, which combined peak intensity and duration across a syllable, accurately described subjects' stress judgments. However, as Sluijter et al. (1997) notice, this measure is confounded with syllable duration. They argue that spectral tilt (the degree of intensity in the higher frequency regions of the spectrum in relation to intensity in the lower frequency regions) better captures the variations in loudness related to word stress than other measures of intensity (Sluijter and van Heuven, 1996a; Sluijter et al., 1997) by showing that duration and spectral tilt were strong correlates of the stress contrast in Dutch, while overall intensity was a correlate of accent rather than stress. On the perceptual side, they showed that a combination of longer duration and flatter spectral tilt was more effective than longer durations and louder overall intensities in differentiating stressed from unstressed syllables in unaccented contexts. Similar results were found for American English (Sluijter, 1995; Sluijter and van Heuven, 1996b). They concluded that spectral tilt was a strong correlate of word stress because it reflected the greater articulatory effort associated with the production of more prominent syllables.

However, Campbell and Beckman (1997) could not replicate Sluijter's results for American English. They showed that spectral tilt did not differentiate stressed from unstressed syllables in the absence of a pitch accent. They also observed that their four subjects varied in how reliably they used other cues, mainly duration. Campbell and Beckman (1997) attributed these apparently conflicting results between English and Dutch to the different use that these languages make of vowel reduction: "Dutch differs from English in having relatively fewer words in which unstressed syllables are reduced" (p. 70). In fact, Sluijter et al. (1997) examined the perception of spectral tilt in relation to stress only in [a], which in Dutch reduces to [a] by lowering F1 and F2 and raising F3 (Sluijter and van Heuven, 1996a, p. 2481), and consequently they could not disentangle the effects of stress in spectral tilt from those of vowel reduction.

Studies of other languages and English dialects also do not reach a consensus on the role of spectral tilt as a cue to word stress or to accent. Crosswhite (2003) shows that phones-DB, an acoustic measure linked to spectral tilt, was significantly affected by stress in Polish, Macedonian and Bulgarian. Kochanski and colleagues (2005) demonstrate that variations in loudness and duration, but not in spectral tilt or F0, are reliable correlates of word stress across seven dialects of British English. In Spanish, where in contrast to Dutch or English, vowels maintain their qualities across stress contexts, listeners did not use variations in spectral tilt to perceive word stress in unaccented sentences (Ortega-Llebaria, Prieto, and Vanrell, 2008). In Swedish, Heldner found that spectral tilt was a strong correlate of focal accent (Heldner, 2001b, 2003). However, he did not demonstrate that it was also a good perceptual cue (Heldner, 2001a) since focal accents in which the spectral tilt had been manipulated were not consistently perceived as more prominent by listeners.

In summary, fifty years of research on stress accent languages provided cumulative evidence that duration is a consistent correlate of stress at the word level, while F0 excursions covary with stressed syllables at the sentence level. Yet, the relation of amplitude changes in the speech signal

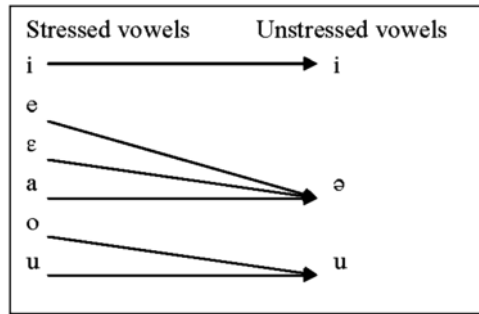
to word stress is still controversial. Despite psychoacoustic research showing that small changes in amplitude are perceptible (Sorin, 1981), no single amplitude measure captures effects on stress judgments across languages. Spectral tilt, for instance, is a consistent correlate of word stress in Dutch, but similar results could not be replicated for Spanish, while for English, the role of spectral tilt in marking word stress remains controversial.

These conflicting results have been attributed to the possible effect of vowel reduction on spectral tilt. The tilt of a vowel spectrum relates to an increase or decrease in the relative amplitude of the higher or lower formants, which can be caused by some change in phonation type, as described in Sluijter et al. (1997, p. 503): “We assume that these differences in the higher parts of the spectrum are caused by a difference in the shape of the glottal waveform, due to an increase in vocal effort when producing stressed syllables, and therefore a reflection of effort, and are perceived in terms of greater loudness.”

In addition to phonation type, changes in the relative amplitude of higher or lower formants are related to the differences in the formant frequencies that make up the quality of a vowel (Fulop, Kari, and Ladefoged, 1998, p. 89), like those formant frequency differences present in the alternation between a stressed open [a] and an unstressed mid-central [ə] in Central Catalan (Herrick, 2003; Recasens, 1986; Recasens and Espinosa, 2006). Yet, to our knowledge, no single study on spectral tilt has controlled for the effects of vowel quality, and since changes in vowel quality tend to covary with stress, the effects of stress on spectral tilt have not been disentangled from those of vowel quality.

The relevance of Castilian Spanish and Central Catalan to the general literature on the correlates of stress is that these two languages offer the appropriate comparison to disentangle the effects of stress on spectral tilt from those of vowel quality by comparing vowels that maintain their qualities across stress contexts with those that change it. In Central Catalan, certain vowels change systematically their qualities in unstressed syllables. There are seven vowels in stressed contexts, that is, [i, e, ε, a, ə, o, u], and three in unstressed contexts, that is, [i, ə, u] (Herrick, 2003; Mascaró, 2002; Recasens, 1986; Recasens and Espinosa, 2006). In unstressed positions the vowels [i] and [u] are unreduced, surfacing as [i] and [u] respectively (*vi* [ˈbi] ‘wine’ > *vinet* [biˈnet] ‘wine.dim’, *suc* [ˈsuk] ‘juice’ > *suquet* [suˈket] ‘juice.dim’) (Recasens, 1986, p. 131). However, the vowels [e, ε, ə] are merged in unstressed positions and surface as schwa [ə] (*peix* [ˈpeʃ] ‘fish’ > *peixet* [pəˈʃet] ‘fish.dim’, *nèt* [ˈnet] ‘nephew’ > *netet* [nəˈtet] ‘nephew.dim’, *pal* [ˈpal] ‘stick’ > *palet* [pəˈlet] ‘stick.dim’), and the vowels [ɔ, o] merge and surface as [u] (*boca* [ˈbokə] ‘mouth’ > *boqueta* [buˈketə] ‘mouth.dim’, *col* [ˈkɔl] ‘cabbage’ > *coleta* [kuˈletə] ‘cabbage.dim’) (Herrick, 2003). The graph in Figure 1 illustrates the vowel inventory in stressed and unstressed positions in Central Catalan.

Therefore, by examining the correlates of stress in vowels that maintain the same quality in both unstressed and stressed contexts, like [i], and in vowels like [a], which alternates with a schwa in unstressed contexts (mainly by lowering F1 about 200 Hz; Recasens, 1986; Recasens and Espinosa, 2006), we control for the effect of vowel quality changes on the stress correlates within the same language. In contrast, speakers of Castilian Spanish use the same five vowel phonemes [i, e, a, o, u] in both stressed and unstressed syllables (Navarro-Tomás, 1914, 1964; Quilis and Esgueva, 1983). By comparing stress correlates of similar vowels in Spanish and Catalan, we control for the effect of vowel quality changes on stress correlates across these languages. Moreover, by examining these correlates of stress in accented and unaccented sentences, we also control for the effect of pitch accents in the expression of the stress contrast. Thus the objective of this article is to examine amplitude and duration correlates of stress in Catalan and Spanish while controlling for pitch accents and the formant frequency differences related to vowel reduction. In particular, the effects of stress on spectral tilt will be examined in vowels with different qualities in order to disentangle the effects of stress on spectral tilt from those of vowel reduction.



**Figure 1.** Inventory of vowels in Central Catalan. Arrows indicate reduction patterns

The article is organized as follows. In the next section, we describe the methodology used for the production experiment. In the results section, we first examine the pitch accents and vowel reduction patterns in our data in order to ensure that we are in fact controlling for the possible effects of accent and vowel quality on stress. Then, we investigate whether the stress contrast is expressed in unaccented contexts as well as in accented contexts by entering all measurements together, that is, duration, overall intensity, spectral tilt and vowel reduction in a Linear Discriminant Analysis that classifies each target vowel as stressed or unstressed. Afterwards, we examine separately the cues of duration, overall intensity and spectral tilt in relation to stress across intonation contexts and vowels. In the discussion section, we compare our results with results reported for other languages.

## 2 Methodology

### 2.1 Materials

Since the goal of this experiment is to examine the acoustic correlates of the stress contrast, it is crucial to employ test materials where (1) stress has a high functional load, and (2) the possible effects of accent and vowel reduction on stress are controlled for. In Spanish and Catalan, stress has a high functional load in the verbal paradigm, where it is very common to find minimal pairs based on stress. Thus, our materials included 15 minimal pairs, like ‘*determino/determinó*’ (‘I establish/(s)he established’), which have cognates in Catalan and in Spanish (see Appendix A). Each verb had four syllables and ended either in *-nimar*, like *desanimar* (‘to discourage’), or in *-minar*, like *determinar* (‘to establish’).<sup>1</sup>

The target verbal forms had either a paroxytone stress in the present tense (i.e., *desanimo* ‘I discourage’, *determino* ‘I establish’), or an oxytone stress in the past tense (i.e., *desanimó* ‘(s)he discouraged’ and *determinó* ‘(s)he established’). In this way, we were able to establish paradigmatic comparisons between syllables that had the same phonemic content and that differed only in degree of prominence. For example, stressed *mó* in *desanimó* was contrasted with unstressed *mo* in *desanimo*, and stressed *ní* in *desanimó* with unstressed *ni* in *desanimó*. These verbs also allowed for syntagmatic comparisons, where, for example, the duration ratio of *mi/no* from the paroxytone verbs was compared to the duration ratio *mi/no* from oxytone verbs in both accented and unaccented contexts.

Covariation between stress and accent is common in stress accent languages like Spanish and Catalan because pitch accents are usually associated with the stressed syllable of every content word

in declarative sentences. This covariation can be disentangled if the stress contrast is examined not only in accented but also in deaccented words. In Spanish and Catalan, we find deaccented words in reporting clauses. Navarro-Tomás (1948, p. 116) defines reporting clauses as the “insertion of an additional element, which has its own characteristics and whose melody is independent of that of the main clause”.<sup>2</sup> In contrast to the main clause, reporting clauses in Spanish are produced in a monotone low pitch melody which has a slight descending trajectory (see Figure 2b). This characteristic monotone intonation does not include any noticeable pitch excursions (Navarro-Tomás, 1914, 1948; and see Astruc, 2005, for similar descriptions of reporting clauses in Catalan).

Thus, the target verbal forms in our test materials were embedded in utterance fragments that would elicit either a declarative intonation or the monotone intonation of reporting clauses. As illustrated in Figure 2a, the utterance *Determino la masa* ‘I establish its mass’ is spoken with broad-focus declarative intonation. The target verb *determino* bears a pitch accent which starts rising at the beginning of the stressed syllable *mi* and reaches its peak in the following post-tonic syllable *no*.<sup>3</sup> Thus, the stressed syllable *mi* is spoken with both *stress and accent*. In 2b, however, *Determino complacida* ‘(S)he is pleased to establish it’ is spoken with the typical flat F0 contour of reporting clauses. Even though syllables *mi* and *no* in *determino* differ with regards to stress, both syllables are produced with a flat F0 contour – as measurements in the following sections will corroborate – indicating that they are deaccented. Consequently, in reporting clauses, *stressed* syllables, like syllable *mi*, are *deaccented*.

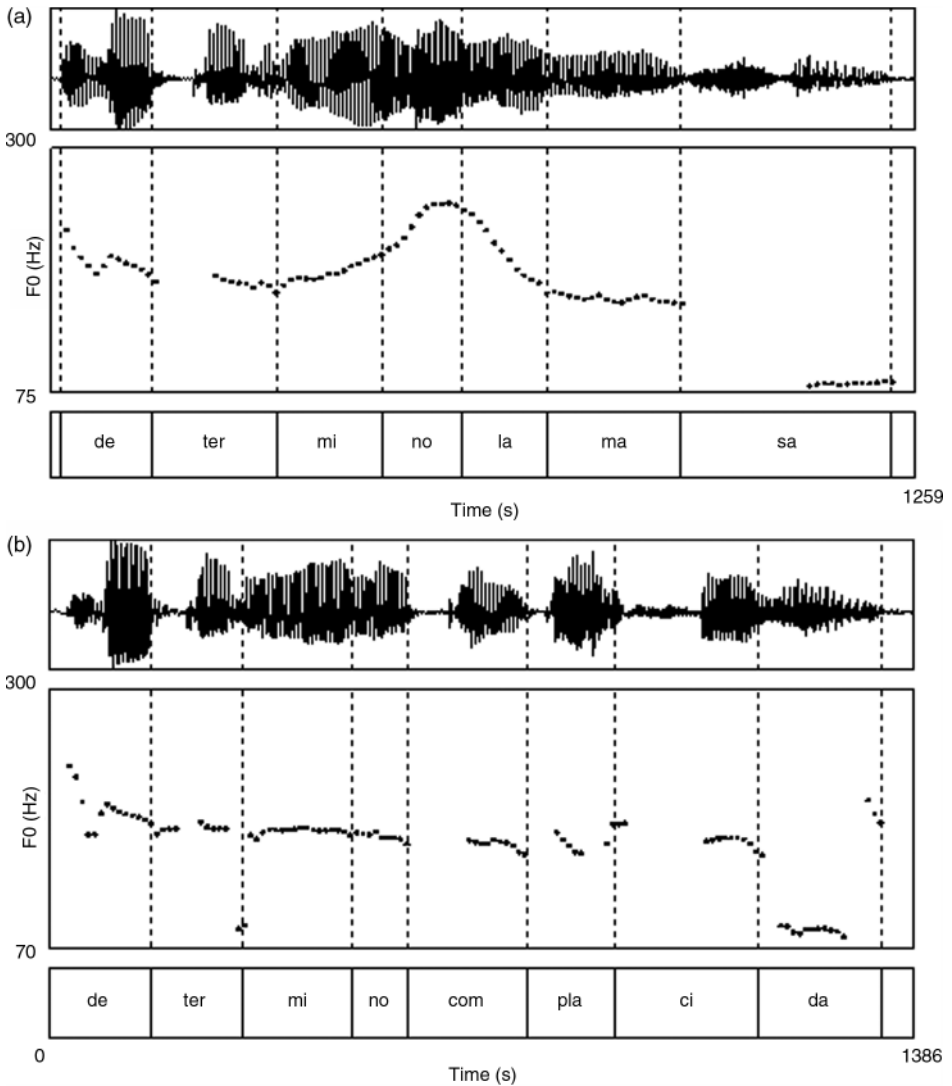
Embedding each verb in both a declarative sentence and a reporting clause yields the combinations shown in Table 1, which allow us to control for the covariation between stress and accent by examining the [+stress] vs. [–stress] contrast in accented and in deaccented words. In declarative sentences, the [+stress] vs. [–stress] contrast is produced along with an accent contrast. For example, stressed *mi* in sentence (a) is also accented, while unstressed *mi* in sentence (b) is deaccented. In reporting clauses, however, stressed *mi* in sentence (c) and unstressed *mi* in sentence (d) are always deaccented.

Many languages accompany the stress contrast with changes in vowel quality: stressed vowels are realized with more peripheral qualities while unstressed vowels become reduced, that is, are more central. The effect of vowel reduction on stress is controlled for in our test materials by comparing Central Catalan, which has phonological vowel reduction related to stress (Mascaró, 2002; Recasens, 1986; Recasens and Espinosa, 2006), with Castilian Spanish, which has not (Quilis, 1981; Quilis and Esgueva, 1983, among others). In Central Catalan, stressed [a] alternates with [ə] in unstressed contexts while [i] maintains the same vowel quality in stressed and unstressed contexts. Then, in the *-mina* (or *-nima*) endings of our target verbs, the quality of the penultimate vowel [i] will remain unchanged regardless of stress, while the last vowel will be realized as [a] in the oxytone verbs, like *deter*[mi'na], or it will be produced with a schwa in the paroxytone verbs, as in *deter*['minə]. Therefore, verb pairs like *determina* and *determinə* will differ not only in stress but also in vowel quality.

Unlike Catalan, there is no phonological vowel reduction in Castilian Spanish. Speakers produce the same five vocalic phonemes in stressed and unstressed positions. Therefore, the vowel qualities of [i] and [o] in the verb endings *-mino* and *-minó* (or *-nimo* and *-nimó*) will remain unchanged regardless of their bearing a stress or not. Consequently, *determino* and *determinó* will be true minimal pairs in Castilian Spanish because they only contrast in stress and not in vowel quality.

Examples of the sentences given to the 20 speakers are included in Appendix B. The data result in a total of 2,400 syllabic tokens: 2 syllabic positions (ultimate and penultimate) × 2 utterance-types (declarative and reporting clause) × 15 verbs × 2 verb tenses × 10 subjects × 2 languages.





**Figure 2.** F0 trajectories in declarative sentences (2a) and reporting clauses (2b). Oscillogram with F0 track and segmentation tier of the declarative utterance *Determino la masa* ‘I establish the mass’ in 2a and the reporting clause *Determino complacida* ‘I establish happily’ in 2b. Note the F0 trajectory on the stressed syllable *mi*. In 2a, F0 starts rising on *mi* and reaches its peak in *no*, whereas F0 shows a flat trajectory in 2b

### 2.2 Experimental procedure

Subjects were presented with 30 cards which contained a verb in infinitival form, a context, and two questions with their answers. As shown in the example in Appendix C, the first answer was to be spoken in a declarative intonation and the second with the flat intonation of extrasentential reporting clauses. The experimenter, either the first or the second author, and the subject were sitting next to each other, in front of a table with the pile of 30 cards and a standing microphone placed close to the subject. After shuffling the 30 cards, the experimenter chose the card on top of the pile, placed it on

**Table 1.** Syllable *mi* in declarative and reporting sentences

Stress and accent contexts for <b>mi</b>	Declarative sentences: [+accented] in [+stress] contexts	Reporting clauses: [-accented] in [+stress] and [-stress] contexts
[+stress] paroxytone verbs	(a) <i>Deter<u>mi</u>no la masa</i> (S)he establishes its mass	(c) <i>La masa del átomo es medible – deter<u>mi</u>no complacida</i> The atom's mass is measurable – I confirm satisfied
[-stress] oxytone verbs	(b) <i>Deter<u>mi</u>nó la masa</i> (S)he established its mass	(d) <i>La masa del átomo es medible – deter<u>mi</u>nó complacida</i> The atom's mass is measurable – (s)he confirmed satisfied

the table between her and the subject and read the context out to the speaker. Once the subject indicated he/she understood the context and the task, the experimenter read out one of the two questions on the card, and the subject read out the corresponding answer with an appropriate intonation. Afterwards, the remaining question was read and the subject read out the appropriate answer. If the experimenter thought that the utterance was mispronounced, for example, there were hesitations, coughing, inappropriate intonation or pauses, the experimenter gave feedback to the subject and he/she was asked to repeat the sentence. The same process was repeated with each one of the cards.

Speakers were recorded individually in a quiet room, using an omnidirectional condenser microphone. Speech samples were recorded digitally at 32,000 Hz, 16-bit.

### 2.3 Subjects

Ten native speakers of Castilian Spanish (7 females and 3 males) and ten native speakers of Central Catalan (10 females) participated in the experiment. Their ages ranged from 20 to 32 years old. All subjects had earned university degrees and spoke an educated variety of their native tongue which they used regularly with their parents and siblings. The ten Catalan subjects had learnt Spanish as a second language in school at age five or later while only five of the Spanish subjects also knew Catalan as a second language, which they learnt in secondary school, that is, at age ten or later. No subject reported having speech or hearing problems.

### 2.4 Measurements

**2.4.1 Fundamental frequency.** We take the widely accepted view that F0 movements are correlates of pitch accents (Ladd, 1996; Pierrehumbert, 1980). In order to ensure that pitch accents appeared only in declarative sentences and not in reporting clauses, we manually marked each syllabic boundary of the verb endings, that is, *mino* or *nimo* in Spanish and *mina* or *nima* in Catalan, yielding three marks, one at the beginning of the first syllable (*mi* or *ni*), a second mark between the first and second syllable (between *mi* and *no* or *ni* and *mo* in Spanish, and between *mi* and *na* or *ni* and *ma* in Catalan) and a third one at the end of the last syllable. A Praat (Boersma and Weenink, 2005) script extracted the F0 value in Hz at each marked point.

**2.4.2 Duration.** Each segment of the verb endings *mino* and *nimo* (for Spanish) and *mina* and *nima* (for Catalan) was marked by hand by the two authors. Nasals sounds are characterized by an abrupt



change in intensity from the preceding vowel (Stevens, 2000). Based on the intensity changes visible in the spectrogram, a portion of the vowel and its following nasal sound were closely examined. The beginning of the nasal was marked at the zero crossing of the first cycle in the waveform that showed a sudden decrease in peak intensity. F2 movements were used to locate vowel transitions (Stevens, 2000). The beginning of vowel transition was marked at the start of the ascending movement departing from the steady F2 of the nasal sound. The end of the last vowel was located at the end of its descending F2 trajectory before starting the next consonant. Formant transitions between vowels and consonants were included in the vowel segments. The tokens that were considered difficult to label by one of the authors, which represented less than 3% of the total, were transcribed by both authors until they reached a consensus. A script extracted the duration in milliseconds of each segment and each syllable.

**2.4.3 Vowel quality.** Formant measurements were based on the stable part of the vowel as shown by a steady F2. The average frequency in Hertz for F1, F2 and F3 across the vowel was calculated by finding formants from the Linear Prediction Coding (LPC) coefficients, a common technique for finding broad spectral peaks (Johnson, 2003, p. 41), which were computed with the Burg algorithm. Window length was set to 25 ms, and frequency ceiling values were set to 5,000 Hz for male and 5,500 Hz for female speakers. The formant values obtained with LPC were checked for plausibility.

**2.4.4 Intensity.** We measured intensity as both overall intensity and spectral tilt. While overall intensity estimates the loudness of a sound by averaging intensities across all frequencies of the spectrum, spectral tilt measures the intensity differences between the lower and the higher frequency regions of the spectrum. Since the increase in the relative amplitudes of the higher formants may be due to differences in the frequencies of the formants, we calculated two measures of spectral tilt. A1-A2 calculates spectral tilt as the intensity differences between the first and the second formants of the target vowel.<sup>4</sup> These measurements are an approximation of spectral tilt at the output sound.

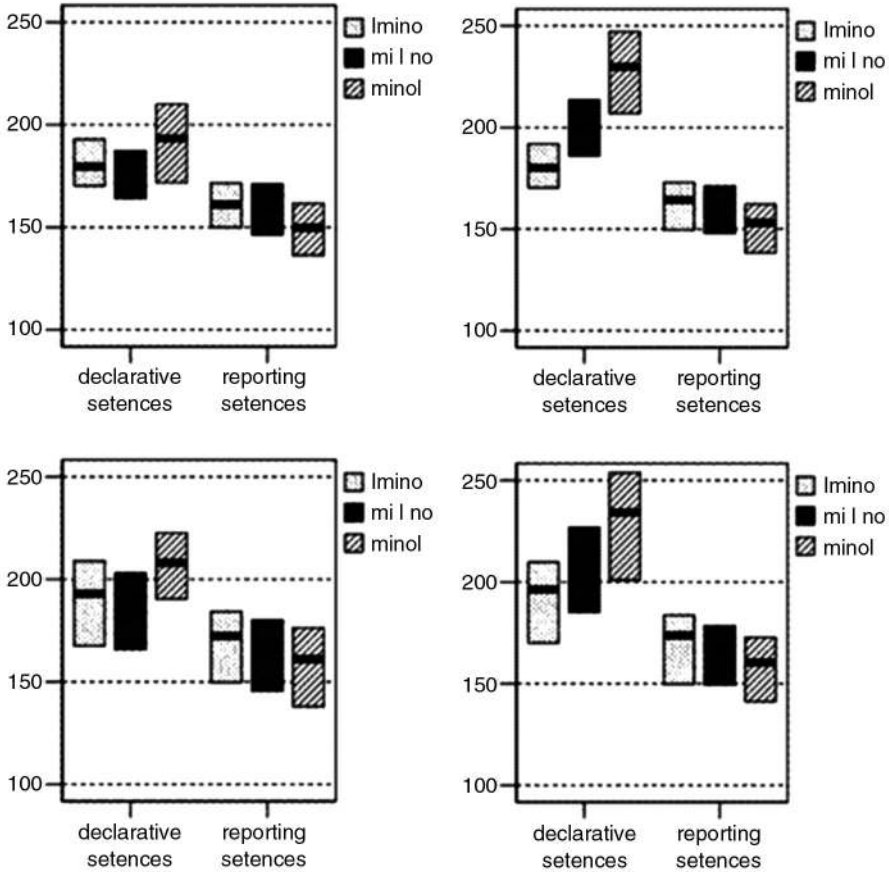
Second, we used the Fulop, Kari, and Ladefoged (1998) equations in order to obtain \*A1-A2 measurements of spectral tilt. In Fulop and colleagues' words, this measurement of spectral tilt "can be regarded as a measure of relative formant amplitude, corrected so as to take into account differences due to formant frequencies" (1998, p. 91). Thus, in our data, \*A1-A2 differs from A1-A2 in that it estimates the spectral tilt of the measured vowel correcting for the formant frequency differences related to vowel reduction, while A1-A2 does not include these corrections.

## 3 Results

### 3.1 Controlling for pitch accents and vowel quality

**3.1.1 Pitch accents and sentence type.** In order to check whether word stress was consistently cued by a pitch accent only in declarative sentences and not in reporting clauses, we measured the F0 values at each syllabic boundary of the *mi-no* verb endings (*mi-na* in Catalan) and compared these values across sentence types, verb tenses and languages. Pitch accents were realized as an upward pitch excursion that started at the beginning of each stressed syllable. For example, when the verb is oxytonic, that is, *determinó*, the rising pitch trajectory started at syllable *no*, and in paroxytonic verbs, that is, *determino*, at syllable *mi*. Since these were prenuclear pitch accents, the peak of this rising accent was realized in the post-tonic syllable.

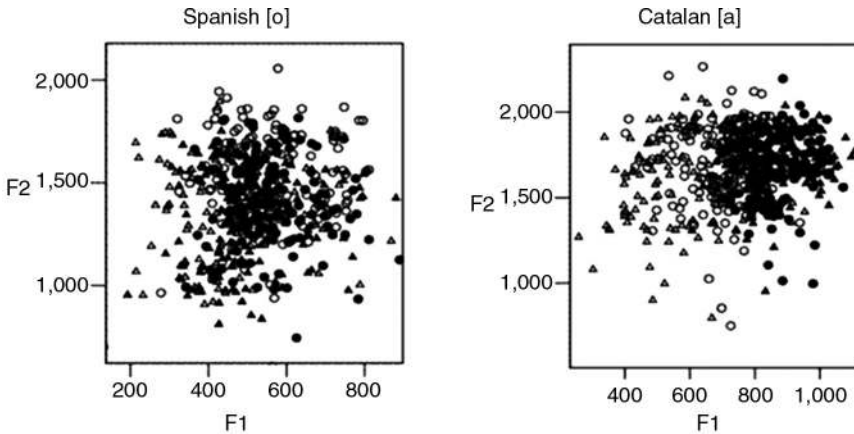
Each graph in Figure 3 portrays the F0 values at the beginning of *mi* in the *mino* ending (dotted box plots), at the syllable boundary between *mi* and *no* (dark box plots), and at the end of *mino* (box



**Figure 3.** Pitch accents in declarative sentences and reporting clauses. Box plots represent the 25th, 50th and 75th percentiles of the F0 value distributions at each syllable boundary of the verb endings, i.e., 'lmino, at the beginning of 'mi', mi l no, between 'mi' and 'no', and mino! at the end of 'no'. Verbs were stressed either in the last syllable (oxytone stress, graphs on the left) or in the penultimate syllable (paroxytone stress, graphs on the right), and were spoken in sentences with declarative intonation or in reporting clauses by Catalan (the two graphs at the top) and Spanish speakers (the two graphs at the bottom). The ascendant trajectories of pitch accents appear only in declarative sentences

plots with stripes) in declarative and reporting sentences in oxytone (two graphs on the left) and paroxytone (two graphs on the right) verbs spoken in Catalan (two graphs at the top) and Spanish (two graphs at the bottom). As expected, the graphs show that there are no rising pitch trends in reporting sentences (the three box plots on the right in each case), while in declarative sentences (on the left) pitch starts increasing on syllable *no* for oxytone verbs and on the syllable *mi* for paroxytone verbs.

Multiple Comparisons tests with the Bonferroni Adjustment compared the F0 values at consecutive syllable boundaries for each sentence type, verb tense and language. There are no significant differences in reporting sentences,  $p > .05$  in all comparisons, demonstrating that speakers produced no pitch accents in this sentence type. However, these differences were highly significant in declarative sentences. As expected, F0 started rising in stressed syllables since it had a



**Figure 4.** Vowel quality of stressed and unstressed vowels. The two graphs show the values of  $F1 \times F2$  in Hz for Spanish [o] and Catalan [a] in stressed (filled shapes) and unstressed (unfilled shapes) syllables in declarative sentences (triangles) and reporting clauses (circles). While there are no significant F1 or F2 differences between stressed and unstressed [o], stressed [a] had a higher F1 than its unstressed counterpart, i.e., [ə]

significantly higher value at the end of the stressed syllable than at its beginning (in paroxytone words like *determino*, Catalan:  $p < .0001$ , Spanish:  $p = .002$ ; in oxytone words like *determinó*,  $p < .0001$  in Catalan and Spanish). This pitch rise continued in the post-tonic syllable of paroxytonic verbs, where the pitch accent reaches its peak, that is, *no* in *determino*,  $p < .0001$  in Catalan and Spanish. In the pre-tonic syllables of oxytone words, that is, *mi* in *determinó*, there were no significant F0 differences between beginning and end,  $p > .05$  in both languages.

In summary, declarative sentences were spoken with pitch accents while reporting clauses were not. Thus, our data are appropriate to control for the effect of pitch accents in the stress contrast by comparing stressed and unstressed syllables in declarative sentences, where there is covariation between stress and accent, with those in reporting clauses, where all syllables are unaccented.

**3.1.2 Vowel quality in Catalan and Spanish.** Spanish and Catalan differ greatly in the way they use vowel reduction to mark the stress contrast. As explained in the introduction, Central Catalan has a process of vowel reduction that affects mainly vowel quality (Recasens, 1986; Recasens and Espinosa, 2006). However, in Castilian Spanish there is no stress-governed vowel reduction (Quilis, 1981; Quilis and Esgueva, 1983). Therefore, we expect that Catalan and Spanish speakers will maintain the quality of vowel [i] unaltered in the endings of target verbs regardless of whether or not they are stressed, that is, stressed [i] in *determino* has the same vowel quality as unstressed [i] in *determinó* (or *determinina* and *determinà* in Catalan). Likewise, Spanish speakers are expected not to change the vowel quality of [o] in *determino* with respect to that in *determinó*. However, Catalan speakers will produce the stressed [a] of *determinà* as a full vowel while in paroxytone verbs like *determinina*, they are expected to pronounce this last vowel as a schwa. These two vowels differ mainly in their F1 values with the F1 of [a] about 200 Hz higher than that of the unstressed [ə] (Recasens, 1986; Recasens and Espinosa, 2006).

The two graphs in Figure 4 show the values of  $F1 \times F2$  in Hz for Spanish [o] and Catalan [a] in stressed (filled shapes) and unstressed (unfilled shapes) syllables across declarative sentences (triangles) and reporting clauses (circles). While in Catalan unstressed [a] has a lower F1 than

stressed [a] in both declarative sentences and reporting clauses (graph on the right), no differences were found between stressed and unstressed [o] in Spanish (graph on the left).

Two Repeated Measures ANOVAs were performed across subjects ( $F1$ ) and items ( $F2$ ), one on the first vowel of the verb ending, that is, [i], and a second ANOVA on the second vowel, that is, [a] or [o] (according to the language). Each ANOVA had the within-subject factors of stress ([+stress] vs. [-stress]) and sentence type (declarative sentences vs. reporting clauses) and the between-subject factor of language. The ANOVA was performed on the difference between the  $F2$  and  $F1$  values of the steady part of the vowel. Results corroborate the expected trends. There are no significant differences for vowel [i], indicating that this vowel does not change its quality in relation to either stress or sentence type in either language. The only significant results are for vowels [o] and [a] in relation to stress,  $F1(1,18) = 58.322, p < .0001, F2(1,28) = 105.958, p < .0001$ , and the interaction stress\*language,  $F1(1,18) = 6.073, p = .024, F2(1,28) = 11.162, p = .002$ . Post-hoc tests with the Bonferroni Adjustment compared the means of stressed [o] or [a] with their unstressed counterparts within and across intonations. While there were no significant results for Spanish [o], Catalan stressed [a] was significantly different from unstressed [ə] within and across intonations,  $F1 p < .05, F2 p < .05$ , confirming, as shown in Figure 4, that stress had an effect on the vowel quality of Catalan [a], but not on Spanish [o].

To sum up, the vowel reduction patterns in our data were as expected. The vowels of the *mino* verb ending in Spanish maintain the same quality across stress contexts. However, in the *mina* ending of Catalan verbs, only [i] maintained the same quality across stress contexts, while stressed [a] alternated with unstressed [ə].

### 3.2 The stress contrast in reporting sentences: Linear Discriminant Analyses

Most prior studies have examined stress in Catalan and Spanish only in declarative sentences (for Spanish, see Canellada and Kuhlman-Madsen, 1987; Contreras, 1963, 1964; Enríquez et al., 1989; Gili Gaya, 1975; Llisterri et al., 2003; Navarro-Tomás, 1914, 1964; Quilis, 1971; and Solé, 1984 among others; for Catalan, see Barnils, 1933; and Recasens, 1986), and consequently, they find  $F0$  to be the main correlate of word stress. For example, Llisterri et al. (2003) found that stress was perceived in syllables that had high  $F0$  values together with longer syllabic durations or louder intensities, but not by any one cue alone, implying that the stress contrast may be lost in deaccented contexts. In order to examine whether speakers' productions express the stress contrast not only in declarative sentences but also in reporting clauses, we performed two sets of Linear Discriminant Analyses (LDAs). In the first set of LDAs, z-scores on measurements of duration, overall intensity, spectral tilt and vowel quality are entered together as independent variables for vowels [a] and [o] (according to the language), and [i]. These measurements were based on paradigmatic comparisons between a stressed vowel, for example, [i] in paroxytone *determino*, and its unstressed counterpart, for example, [i] in oxytone *determinó*. Percentages of vowels correctly classified as stressed or unstressed in declarative sentences and reporting clauses were clearly above chance in all speakers (Table 2), indicating that the stress contrast is expressed in the speakers' productions. Multiple Comparisons on the speakers' classification percentages show that there are no significant differences between declarative and reporting sentences in any language and vowel (Catalan: [a]  $p = 1.000$ , [i]  $p = 1.000$ ; Spanish: [o]  $p = 1.000$ , [i]  $p = .279$ ) confirming that the stress contrast is expressed not only in declarative sentences but also in reporting clauses in all three vowels regardless of the reduction patterns.

Second, LDAs were also performed on duration ratios, overall intensity differences, spectral tilt differences, and vowel quality ratios, that is,  $F2-F1$  in vowel 2 /  $F2-F1$  in vowel 1, between the second and first vowel of the verb endings (Table 3), and Multiple Comparisons show no

**Table 2.** Percentages of syllables correctly identified as stressed or unstressed when duration, overall intensity, spectral tilt and vowel reduction were entered as independent variables. Syllables are specified according to vowel and language

	Spanish [o]		Spanish [i]		Catalan [a]		Catalan [i]	
	Declarat.	Report.	Declarat.	Report.	Declarat.	Report.	Declarat.	Report.
Mean	81.41	78.24	78.56	69.58	93.64	95.01	78.95	72
St. dev.	8.4	10.1	8.3	8.6	6.0	3.3	13.5	8.2

**Table 3.** Percentages of words correctly classified as oxytone or paroxytone when duration ratios, overall intensity differences, spectral tilt differences and vowel reduction ratios were entered as independent variables

	Spanish		Catalan	
	Declaratives	Reporting	Declaratives	Reporting
Mean	83.01	77.86	89.26	86
St. dev.	7.08	12.48	5.17	10.02

significant differences between declarative and reporting clauses within languages (Catalan:  $p = 1.000$ , Spanish:  $p = 1.000$ ) or across languages (Catalan vs. Spanish in declaratives:  $p = 1.000$ , in reporting:  $p = .634$ ). Thus, speakers from both languages also express the stress contrast in their productions across sentence types when stress differences are measured in adjacent syllables.

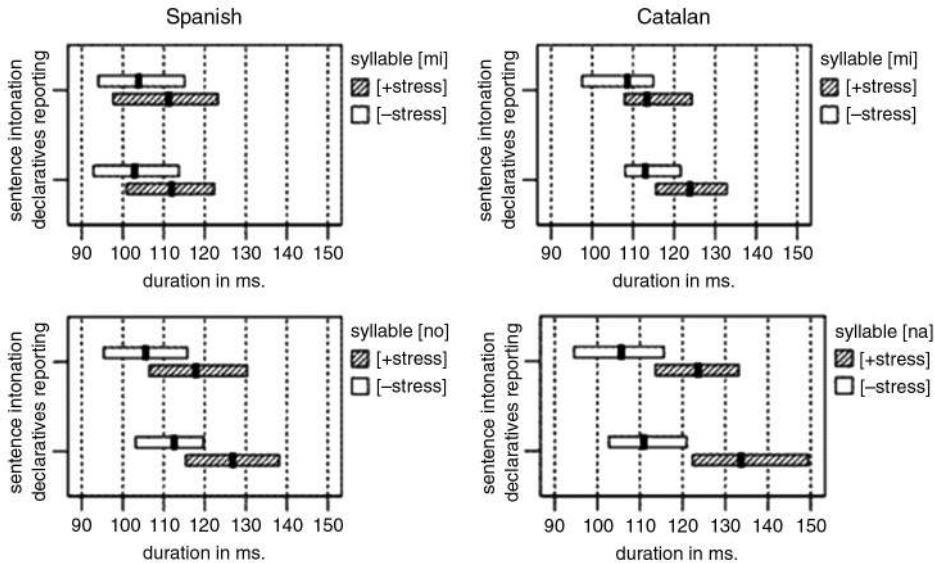
### 3.3 Cues to the stress contrast: syllable duration, overall intensity and spectral tilt

In the following sections we examine each cue to stress. Three-way Repeated Measures ANOVA with the within-subject factors of stress ([+stress] and [-stress]) and sentence type (declarative sentences and reporting clauses) and the between-subject factor of language (Catalan and Spanish) were performed separately on measurements of syllable duration, overall intensity and spectral tilt across subjects ( $F1$ ) and items ( $F2$ ).

**3.3.1 Syllable duration.** Figure 5 displays the mean duration (in ms) of each syllable of verb endings (*mino* in Spanish, *mina* in Catalan). In each graph, the durations of stressed syllables are compared with the duration of their unstressed counterparts in both accented and unaccented tokens. For example, the top graph on the left displays the duration of stressed *mi* as spoken in paroxytone *determino* (box plots with stripes) compared with the duration of unstressed *mi* as in oxytone *determino* (box plots in white) in reporting clauses (the two box plots at the top of the graph) and declarative sentences (the two at the bottom part of the graph). Similar comparisons are depicted for syllable *mi* in Catalan (top graph on the right), for syllable *no* in Spanish (graph at the bottom left), and for syllable *na* in Catalan (at the bottom right).

Figure 5 shows that stressed syllables are consistently longer than their unstressed counterparts across the languages, vowels and sentence types, indicating a clear durational effect of stress. The size of this effect, however, is not homogeneous. Duration differences between stressed and unstressed syllables are larger in Catalan than in Spanish (mean difference in Catalan: 16 ms; in Spanish: 11.1 ms), in the syllable *no* (or *na* in Catalan) than in the syllable *mi* (mean difference



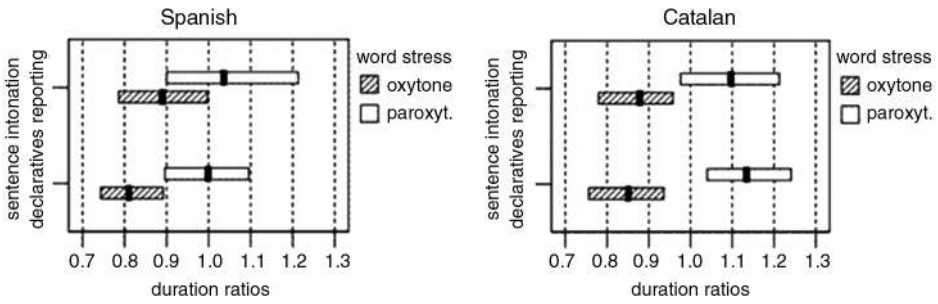


**Figure 5.** Syllable duration in ms of stressed and unstressed syllables. Box plots represent the 25th, 50th and 75th percentiles of the distributions of stressed syllables (shaded rectangles) and unstressed syllables (white rectangles) in reporting clauses and declarative sentences in Spanish (graphs on the left) and Catalan (graphs on the right) for syllables [mi] (top graphs) and [no]/[na] (bottom)

between stressed and unstressed syllables in *no/na*: 18 ms; in *mi*: 8.7 ms), and in declarative sentences compared to reporting clauses (mean difference between stressed and unstressed syllables in declaratives 15.5 ms; in reporting clauses 11 ms). The largest duration differences are between stressed *na*, which contains a full vowel, and unstressed [nə], which contains a schwa (mean difference: 22 ms, while mean differences between stressed and unstressed *no* are 14 ms, *mi* in Catalan: 9.9 ms, and *mi* in Spanish: 8.7 ms). Since the [na]–[nə] alternation is the only context where the stress contrast covaries with changes in vowel quality and open vowels tend to be longer than their mid-central counterparts (Lindblom, 1963; Moon and Lindblom, 1994), then contrasting a full vowel with a reduced vowel amplifies the durational effect of stress.

Results from the three-way Repeated Measures ANOVA performed separately on the durations of syllables *mi* and *na* (*no* for Spanish) across subjects and items show that the main factor of stress was significant for syllable *mi*,  $F(1,18) = 86.379, p < .0001, F(2,28) = 120.705, p < .0001$ , and syllable *no* (or *na* in Catalan),  $F(1,18) = 248.168, p < .0001, F(2,28) = 340.832, p < .0001$ , confirming that stressed syllables are consistently longer than their unstressed counterparts. The interaction ‘stress\*language’ gave significant results only for the last syllable (*na/no*:  $F(1,18) = 13.345, p < .0001, F(2,28) = 17.718, p = .002$ ) verifying that the durational effect of stress in Catalan *na* is larger than that of Spanish *no*, whereas these two languages are no different in the duration differences of *mi*. These results corroborate that stress has a durational effect which is pervasive across vowels and languages, and that the changes in vowel quality present in Catalan amplify the duration differences related to stress.

Results from the ANOVA tests also indicated a significant effect of sentence type: syllables in declarative sentences were consistently longer than those in reporting clauses (*mi*:  $F(1,18) = 15.1198, p = .001; F(2,28) = 28.769, p < .0001$ ; *no*:  $F(1,18) = 49.439, p < .0001; F(2,28) = 27.827, p < .0001$ ). Multiple Comparisons with the Bonferroni Adjustments show that syllables



**Figure 6.** Duration ratios of stressed and unstressed syllables. Box plots represent the 25th, 50th and 75th percentiles of the distributions of oxytone verb endings (shaded rectangles) and paroxytone verb endings (white rectangles) in reporting clauses and declarative sentences in Spanish (graph on the left) and Catalan (graph on the right)

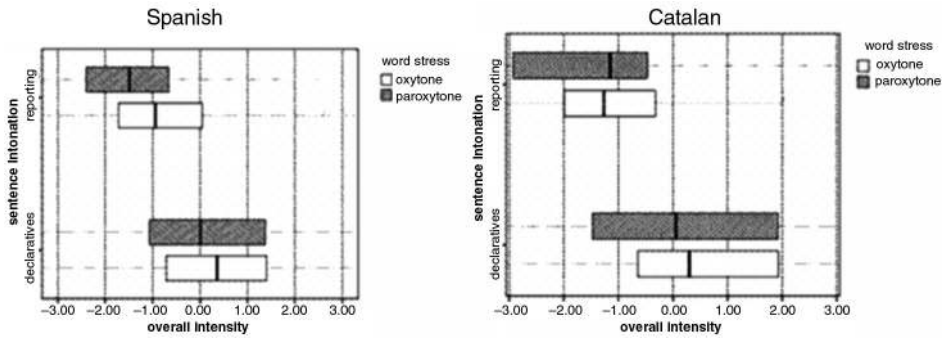
that contrast in accent and sentence type are significantly different. The accented (and stressed) syllables of declarative sentences are longer than the unaccented (and stressed) syllables of reporting clauses, suggesting a possible effect of accent. However, since the syllables that differ only in sentence type and not in accent or stress are also different – for example, unstressed and unaccented syllables in declarative sentences are longer than unstressed and unaccented syllables in reporting clauses (Catalan *mi*:  $p < .0001$ ; Spanish *mi*:  $p < .01$ ; Catalan *na*:  $p < .01$ ; Spanish *no*:  $p < .05$ ) – we interpret this lengthening effect as being related to sentence type rather than to the presence of a pitch accent itself. In fact, Navarro-Tomás (1914, 1948) noticed that reporting sentences tended to be spoken faster than declarative sentences, a pattern that also appeared in our data, that is, target verbs in declarative sentences were on average 12.91 ms longer than in reporting clauses.

In order to control for possible speech rate effects, we analyze patterns of relative duration, that is, we analyze the duration ratio between adjacent syllables. Figure 6 shows duration of *mi* divided by duration of *na* in oxytone and paroxytone verbs. These syntagmatic comparisons show that the ratios change according to stress position in both languages. In oxytonic verbs, ratios are smaller than 1 while in paroxytonic verbs, ratios are larger than 1 or close to 1. This difference is related to the fact that in oxytonic verbs, final syllables are longer than penultimate syllables, while in paroxytonic verbs both syllables have similar durations.

Repeated Measures ANOVAs show stress to be a significant factor,  $F(1,18) = 226.057$ ,  $p < .0001$ ,  $F(1,28) = 1077.747$ ,  $p < .0001$ , confirming the durational effect of stress. The significant ‘stress\*language’ interaction,  $F(1,18) = 9.321$ ,  $p = .007$ ,  $F(1,28) = 44.010$ ,  $p < .0001$ , shows, as displayed in Figure 6, that these ratio differences are larger in Catalan than Spanish. However, Multiple Comparisons with the Bonferroni Adjustment performed within each language and sentence type indicate that ratio differences between oxytone and paroxytone verbs are maintained within sentence type and language (declaratives: Catalan  $p < .0001$ , Spanish  $p < .0001$ ; reporting: Catalan  $p < .0001$ , Spanish  $p < .0001$ ), providing further evidence that the stress contrast works independently from pitch accents and vowel reduction patterns.

In summary, the durational effect of stress consistently appears in speakers’ productions across sentence types and languages confirming that duration is a strong correlate of stress. It is not parasitic on vowel reduction because it takes place in vowels that maintain the same vowel quality across stressed and unstressed syllables. Duration differences between stressed and unstressed vowels were largest in [a], the only vowel that alternates with [ə] in unstressed contexts. Thus, the





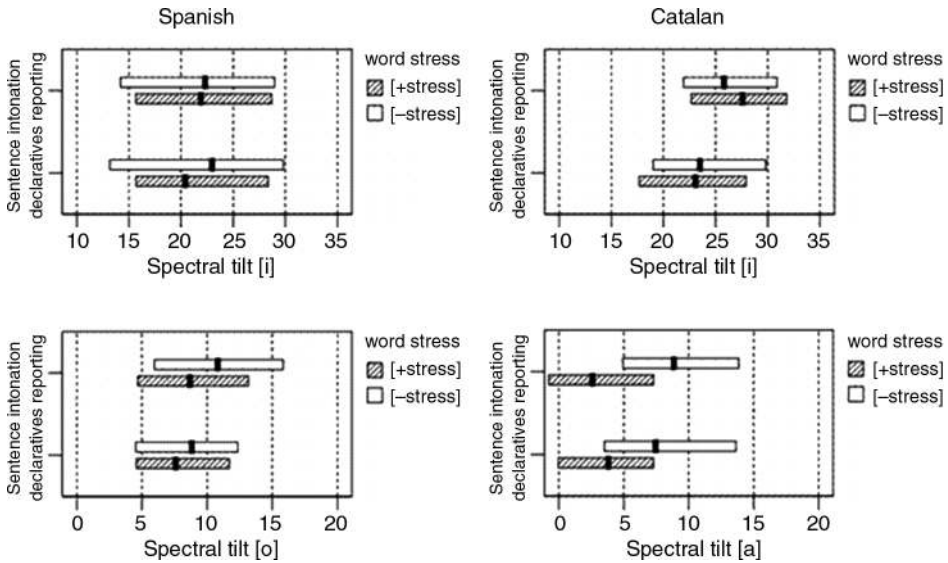
**Figure 7.** Overall intensity differences (dB) between the second and first vowel of the verb endings (*-mino* in Spanish and *-mina* in Catalan). Box plots represent the 25th, 50th and 75th percentiles of the distributions of oxytone verb endings (white rectangles) and paroxytone verb endings (shaded rectangles) in reporting clauses and declarative sentences in Spanish (graph on the left) and Catalan (graph on the right)

vowel quality changes related to vowel reduction do not trigger stress-related durational differences, rather they amplify them.

**3.3.2 Overall intensity measurements.** Since recordings were performed with a standing microphone, speakers' distance to the microphone was not held constant, possibly affecting overall intensity measurements. In order to minimize this effect, intensity was normalized by comparing adjacent syllables: the overall intensity in dBs of vowel [i] was subtracted from the overall intensity of the vowel in its adjacent syllable, which is [a] in Catalan, like in *determina*, and [o] in Spanish, like in *determino*. Figure 7 shows that overall intensity effects are small because differences between vowels do not surpass 3 dB, which makes these differences almost undetectable given that 1 to 2 dB difference is effectively imperceptible under normal hearing conditions. In relation to stress, differences in paroxytone verbs (shaded bar) are larger than in oxytone verbs (white bar) only in Spanish reporting sentences (two bars at the top half of the graph). Sentence type has an effect in both languages. While in declarative sentences (lower bars), differences score around 0 indicating that both vowels were produced with similar intensities, in reporting clauses (two bars at the top half of the graphs) differences are less than 0, showing that the first vowel was produced with greater intensity than the second. This difference relates to the general characteristics of reporting clauses which tend to be spoken with progressively lower intensities and a softer voice than declaratives (Navarro-Tomás, 1914, 1948).

Repeated Measures ANOVAs with the between-subject factors of stress (paroxytonic versus oxytonic verbs) and sentence type (declarative versus reporting clauses), and with the between-subject factor of language (Catalan and Spanish) were performed across subjects and items. There are significant differences for the main factors of stress,  $F(1,18) = 4.883, p = .04$ ,  $F(1,28) = 13.900, p < .0001$ , and sentence type,  $F(1,18) = 32.741, p < .0001$ ,  $F(1,28) = 185.636, p < .0001$ . Multiple Comparisons with the Bonferroni Adjustment confirm that only Spanish speakers use overall intensity to mark stress differences (declaratives:  $p = .046$ , reporting:  $p = .036$ ), and that speakers from both languages produce reporting clauses with lower overall intensities than declarative sentences regardless of stress,  $p < .05$ .

Thus, the effects of overall intensity are almost imperceptible since differences between adjacent syllables do not exceed 3 dB. These small differences relate more consistently to sentence



**Figure 8.** A1-A2 spectral tilt of stressed and unstressed syllables. Box plots represent the 25th, 50th and 75th percentiles of the distributions of stressed syllables (shaded rectangles) and unstressed syllables (white rectangles) in reporting clauses and declarative sentences in Spanish (graphs on the left) and Catalan (graphs on the right). Units on the x-axes are arbitrary

intonation than to stress. Speakers from both languages produce reporting sentences with lower overall intensities than declarative sentences, but only Spanish speakers use overall intensity to mark the stress contrast, especially in reporting sentences, possibly corroborating Navarro-Tomás' *acento de intensidad*.

**3.3.3 Spectral tilt.** Spectral tilt was first measured as the \*A1-A2 difference, which calculates spectral tilt while correcting for the formant frequency differences related to vowel reduction. Three-factor Repeated Measures ANOVAs, that is, sentence type (declarative sentence versus reporting clause), stress ([+stress] versus [-stress]) and the between-subject factor of language (Catalan and Spanish), were performed on each vowel ([i] versus [o] or [a]) across subjects and items. There were no significant main effects or interactions, confirming that when the spectral tilt of [i], [a] and [o] is corrected for formant frequency differences, it remains unchanged across stress contexts and sentence types.

However, when spectral tilt was measured without correcting for formant frequency differences (A1-A2 measurements), there is a clear effect of stress in the spectral tilt of Catalan [a], the only vowel in our test materials that alternates with a schwa in unstressed contexts. The graph at the bottom right in Figure 8 shows that vowel [a] has a flatter spectral tilt than the unstressed schwa and this spectral difference remains constant across sentence types. In contrast, vowels that maintain the same vowel quality across stress contexts, that is, [i] in Catalan (top graph on the right), [i] in Spanish (top graph on the left), and [o] in Spanish (bottom graph on the left), do not show differences in spectral tilt.

The three-way Repeated Measures ANOVAs across subjects and items on A1-A2 measures of spectral tilt show a significant effect of stress on the last vowel of the target verb, that is, [a] or [o],  $F(1,18) = 33.155, p < .0001$ ,  $F(1,28) = 47.141, p < .0001$ , and that this effect is different across

languages (stress\*language:  $F1(1,18) = 16.088, p = .001$ ;  $F2(1,28) = 20.202, p < .0001$ ). No significant effects were found for [i] (stress:  $F1(1,18) = .025, p = .877$ ;  $F2(1,28) = .197, p = .661$ ; stress\*language:  $F1(1,18) = .020, p = .890$ ;  $F2(1,28) = .845, p = .366$ ). Repeated Measures ANOVAs performed separately for [a] and [o] corroborate that stress has a significant effect only on the spectral tilt of Catalan [a],  $F1(1,18) = 31.930, p < .0001$ ,  $F2(1,28) = 44.868, p < .0001$ , and not on the spectral tilt of Spanish [o],  $F1(1,18) = 3.018, p = .116$ ,  $F2(1,28) = 5.007, p = .062$ .

In summary, our data show that when measures of spectral tilt were corrected for formant frequency differences in corresponding vowels, no differences were found between stressed and unstressed vowels. Only when spectral tilt was not corrected for formant frequency differences, stressed [a] had a significantly flatter tilt than its unstressed counterpart [ə]. Vowels that maintained the same quality across stress contexts did not show changes in spectral tilt. Thus, in Catalan, changes in spectral tilt are not a correlate of stress proper but they are a consequence of the formant frequency differences associated with vowel reduction.

## 4 Discussion

The contribution of this article to the literature on the phonetics of stress is to study the duration and amplitude correlates of stress in Catalan and Spanish while controlling for the effects of pitch accents and the formant frequency differences related to vowel reduction. As explained in the introduction, F0 and duration are the least controversial of the stress correlates. In stress accent languages, F0 excursions align with stressed syllables, causing covariation between stress and F0. Consequently, the timing of the F0 movements with respect to the syllable is crucial for determining stress perception. Yet, the stress contrast is also preserved in unaccented sentences indicating that there is a trading relationship between F0 and other stress correlates. These trading effects have been shown for many languages (Beckman and Edwards, 1994; Huss, 1978; Lieberman, 1960; Nakatani and Aston, 1978, for English; Sluijter and van Heuven, 1996a; van Heuven, 2002, for Dutch; Manolescu et al., 2009, for Romanian; Dogil and Williams, 1999, for German; Kastrikani, 2003, for Greek; Ortega-Llebaria and Prieto, 2007, for Spanish) but not for Catalan because stress correlates were previously examined only in accented sentences. In this article, results from the LDAs show that Catalan and Spanish speakers consistently express the stress contrast in unaccented contexts as well as in accented sentences. These results, together with those of previous literature of Catalan and Spanish, provide experimental evidence that the trading between F0 and other stress correlates which characterizes stress accent languages takes place in Catalan and Spanish as well.

There is also ample cross-linguistic evidence of the durational effects of stress. Speakers tend to produce stressed syllables with longer durations than their unstressed counterparts, independently of pitch accents (Beckman and Edwards, 1994; Huss, 1978; Nakatani and Aston, 1978, for English; Ortega-Llebaria, 2006, for Spanish). Our results confirm this general pattern for Catalan and Spanish since stressed syllables were significantly longer than unstressed syllables in both accented and unaccented sentences.

Furthermore, Catalan and Spanish offer the appropriate comparison to examine the durational effect of stress while controlling for vowel quality changes in corresponding vowels. In English and other languages there is covariation between stress and vowel reduction, since most open vowels are stressed while mid-central vowels are unstressed. Moreover, Lindblom (1963) and Moon and Lindblom (1994) showed that there is a monotonic relationship between vowel quality and duration because it takes longer to produce an open and more peripheral vowel than an undershot and mid-central vowel. Consequently, the duration differences between open and mid-central

vowels in English, as those found between the two syllables of *papa* in Beckman and Edwards (1994), may be related to the durational effect of stress as well as to the monotonic relationship between vowel quality and duration.

Our materials allow for the examination of the durational effect of stress while controlling for vowel quality changes in corresponding vowels. As explained in the introduction, there is no vowel reduction in relation to stress in Castilian Spanish, and in Central Catalan, vowel reduction takes place only in part of its vowel inventory. Therefore, we examined the durational effect of stress in vowels that alternate open and more peripheral realizations in stressed contexts with mid-central qualities in unstressed positions, that is, [a]–[ə] in Catalan, as well as in vowels that maintain the same qualities across stress contexts, that is, [i] in Catalan and Spanish and [o] in Spanish. Our results reveal that stress affects duration in all vowels and not only in Catalan [a], showing that duration cues mark the stress contrast independently of changes in vowel quality. In addition, the largest duration differences between stressed and unstressed syllables are found in the [a]–[ə] alternation, showing that the changes in vowel quality between corresponding vowels not only provide a spectral correlate to the stress contrast, but they also have the effect of amplifying the duration differences that were triggered by stress.

In contrast with F0 and duration, intensity correlates of stress are more controversial. Despite psychoacoustic research showing that small changes in amplitude are perceptible (Sorin, 1981), no single amplitude measurement captures effects on stress judgments across languages. Our results agree with those of Turk and Sawusch (1996) for English, in that duration is a more consistent correlate of stress than overall intensity. However, when amplitude changes are measured not as overall intensity, but as an approximation to steady-state perceptual loudness such as Stevens' Mark VII computation (Stevens, 1971), it becomes a consistent correlate of stress in English (Kochanski et al., 2005). Duration differences between stressed and unstressed syllables in Catalan and Spanish are significant across sentence types and vowel categories, while stressed syllables have significantly greater overall intensities only in Spanish reporting clauses, showing that speakers' use of duration is more pervasive than that of overall intensity.

As explained in the introduction, there is no general consensus on how spectral tilt correlates with the stress contrast across languages. In particular, the conflicting results between Dutch (Sluijter, 1995; Sluijter and van Heuven, 1996a, 1996b; Sluijter et al., 1997) and English (Campbell and Beckman, 1997; Sluijter and van Heuven, 1996b) led to the hypothesis that the differences in the vowel reduction patterns of these two languages may have had an effect on their different uses of spectral tilt as a stress correlate. Spectral tilt changes may be caused by the changes in phonation type that Sluijter et al. (1997) relate to the greater articulatory effort of stressed syllables, as well as to the formant frequency differences related to vowel reduction (Fulop et al., 1998). However, to our knowledge, no previous study has addressed the effect of stress on spectral tilt while controlling for vowel reduction.

As for duration, we examined spectral tilt in stressed and unstressed vowels that maintained the same vowel qualities, that is, [i] in Catalan and Spanish, and [o] in Spanish, as well as in Catalan [a], that alternates with [ə] in unstressed syllables. Moreover, Fulop et al. (1998) corrections for formant frequency differences were applied to our measurements of spectral tilt, and statistical tests were run in corrected and uncorrected measurements of spectral tilt. Results indicate that spectral tilt differentiated Catalan stressed [a] from unstressed [ə] in uncorrected measurements of spectral tilt. Vowels that maintained the same qualities across stress contexts did not use spectral tilt to differentiate stressed syllables from their unstressed counterparts, nor did the [a]–[ə] alternation when statistical tests were run on the corrected measurements of spectral tilt. Consequently, the spectral tilt differences in our data are more related to the formant frequency differences between [a] and [ə] than to the stress contrast proper.

These production results are similar to those obtained in Dutch. Sluijter and van Heuven (1996a) found that spectral tilt was a more consistent correlate of stress in [a] than in [o] (1996a, p. 2481). They also found that stressed [a] reduced to unstressed [a] by lowering F1 and F2 and raising F3, while no significant formant frequency differences were found for stressed vs. unstressed [o]. In a perception experiment (Sluijter et al., 1997), they manipulated duration, overall intensity and spectral tilt cues to stress of [a] in *nana*, and found that combinations of duration and spectral tilt better differentiated stressed from unstressed syllables than combinations of duration with overall intensity.

Likewise, perception studies in Catalan (Ortega-Llebaria, Vanrell, and Prieto, 2010) and Spanish (Ortega-Llebaria et al., 2008) show that Catalan listeners, like Dutch listeners, use a combination of duration and spectral tilt in perceiving stress in *mama*. However, Spanish listeners only used duration cues, in spite of performing identical manipulations of duration and spectral tilt in the stimuli. Neither Catalan nor Spanish listeners used spectral tilt in perceiving stress in *mimi*. Thus, spectral tilt was perceived as a cue to the stress contrast only in vowels that alternate a more peripheral realization in stressed contexts with a mid-central quality in unstressed positions.

In sum, spectral tilt is a cue to stress in Dutch, but it relates more to the formant frequency differences related to vowel reduction than to stress proper in Catalan, while it does not carry information about stress in Spanish. In order to disentangle how spectral tilt interacts with vowel reduction to convey a stress contrast, future research needs to examine a set of vowels that alternate open and more peripheral realizations in stressed contexts with mid-central qualities in unstressed positions in a variety of languages.

### Acknowledgements

We would like to thank the audience at the 2005 PaPI Conference, and also Barbara Bullock, Megan Crowhurst, Sónia Frota, José Ignacio Hualde, Scott Myers, Jacqueline Toribio, Daniel Recasens, and Marina Vigário for very useful feedback. We are also grateful to the 20 informants who kindly participated in the production experiment. Thanks to Tim Mills who kindly sent us the script that measures spectral tilt. Special thanks to John Coleman and two anonymous reviewers, whose insightful comments greatly improved the quality of this article. This research has been funded by grant 2005SGR-00753, awarded by the Generalitat de Catalunya, and by grants HUM2006-01758/FILO “Estructura prosódica y adquisición de la prosodia en catalán y español”, CONSOLIDER-INGENIO 2010 “Bilingüismo y Neurociencia Cognitiva CSD2007-00012” awarded by the Spanish Ministerio de Educación y Ciencia, and URA Program awarded by the College of Liberal Arts, University of Texas.

### Notes

- 1 Throughout this article, orthographic transcriptions are in italics, and since orthographic conventions signal only a subset of stressed syllables with a diacritic (an acute accent in Spanish, a grave accent in Catalan), we underline all stressed syllables for clarity. In phonetic transcriptions, stressed syllables are marked with the usual IPA conventions.
- 2 In Spanish, “La intercalación de un elemento incidental, con carácter propio, ajeno a la estructura melódica de la frase en que se encuentra” (Navarro-Tomás, 1948, p. 116).
- 3 Although pitch accents in Spanish align with stressed syllables, only the location of the valleys, e.g., the beginning of the F0 rise, is fairly constant. The location of the peaks, e.g., the end of the F0 rise, varies according to the segmental composition of the stressed syllable as well as the following prosodic context (Prieto, van Santen, and Hirschberg, 1995).
- 4 The script we used was created by Tim Mills and it starts by identifying ranges to look in for maximum LTAS energy – the algorithm looks for the local maximum within 10% of the formant’s average frequency – from which it gets the formant amplitudes and the A1-A2 measure. Then, it calculates the equations used in Fulop et al. (1998) to correct for formant movement, which yields what Fulop and colleagues called the ‘modeled A1-A2’ measure. The \*A1-A2 is the difference between the A1-A2 and the modeled A1-A2.



## References

- Astruc, L. (2005). *The intonation of extra-sentential elements in Catalan and English*. Doctoral Dissertation, University of Cambridge.
- Barnils, P. (1933). Sobre la quantitat de les vocals tòniques. *Escrips de Pere Barnils, Anuari de l'Oficina Romànica de Llengua i Literatura*, VI, 67–78.
- Beckman, M. E. (1986). *Stress and non-stress accent*. Dordrecht: Foris.
- Beckman, M. E., and Edwards, J. (1994). Articulatory evidence for differentiating stress categories. In P. A. Keating (Ed.), *Phonological structure and phonetic form: Papers in laboratory phonology III* (pp. 7–33). Cambridge: Cambridge University Press.
- Boersma, P., and Weenink, D. (2005). *Praat: Doing phonetics by computer* (Version 4.3.01) [Computer program]. Retrieved from <http://www.praat.org/>
- Campbell, N., and Beckman, M. E. (1997). Stress, prominence and spectral tilt. In A. Botinis, G. Kouroupetroglou, and G. Carayiannis (Eds.), *Intonation: Theory, models and applications* (pp. 67–70). ESCA and University of Athens Department of Informatics.
- Canellada, M. J., and Kuhlman-Madsen, J. (1987). *Pronunciación del español. Lengua hablada y literaria*. Madrid: Editorial Castalia.
- Contreras, H. (1963). Sobre el acento en español. *Boletín del Instituto de Filología de la Universidad de Chile*, 15, 223–237.
- Contreras, H. (1964). ¿Tiene el español un acento de intensidad? *Boletín del Instituto de Filología de la Universidad de Chile*, 16, 237–239.
- Crosswhite, K. (2003). Spectral tilt as a cue to stress in Polish, Macedonian and Bulgarian. In M.-J. Solé, D. Recasens, and J. Romero (Eds.), *Proceedings of the XVth International Conference of Phonetic Sciences* (Vol. 2, pp. 767–770). Barcelona: Causal Productions.
- Cutler, A. (2005). Word stress. In D. Pisoni and R. Remez (Eds.), *The handbook of speech perception* (pp. 264–289). Malden, MA: Blackwell.
- Dogil, G., and Williams, B. (1999). The phonetic manifestation of word stress. In H. van der Hulst (Ed.), *Word prosodic systems in the languages of Europe* (pp. 273–311). Berlin: Mouton de Gruyter.
- Enríquez, E. V., Casado, C., and Santos, A. (1989). La percepción del acento en español. *Lingüística Española Actual*, 11, 241–269.
- Fry, D. B. (1955). Duration and intensity as physical correlates of linguistic stress. *Journal of the Acoustical Society of America*, 27(4), 765–768.
- Fry, D. B. (1958). Experiments in the perception of stress. *Language and Speech*, 1, 126–152.
- Fry, D. B. (1964). The dependence of stress judgments on vowel formant structure. In S. Karger Basel (Ed.), *Proceedings of the 5th International Congress in Phonetic Sciences* (pp. 306–311), Münster, Germany.
- Fulop, S. A., Kari, E., and Ladefoged, P. (1998). An acoustic study of the tongue root contrast in Degema vowels. *Phonetica*, 55, 80–98.
- Gili Gaya, S. (1975). *Elementos de fonética general*. Madrid: Gredos.
- Heldner, M. (2001a). Spectral emphasis as a perceptual cue to prominence. *Speech, Music and Hearing: Quarterly Progress and Status Report, THM-QPSR*, Department of Speech, Music and Hearing, Royal Institute of Technology, Stockholm, Sweden, 2, 51–57. Retrieved from <http://www.speech.kth.se/qpsr/tmh/01-2-051-057.pdf>
- Heldner, M. (2001b). Spectral emphasis as an additional source of information in accent detection. *Prosody 2001: ISCA Tutorial and Research Workshop on Prosody in Speech Recognition and Understanding*, Red Bank, NJ, 57–60. Retrieved from <http://www.ling.umu.se/~heldner/papers/prosody2001.pdf>
- Heldner, M. (2003). On the reliability of overall intensity and spectral emphasis as acoustic correlates of focal accents in Swedish. *Journal of Phonetics*, 31, 39–62.
- Herrick, D. (2003). *An acoustical analysis of phonological vowel reduction in six varieties of Catalan*. Doctoral Dissertation in Linguistics, University of California in Santa Cruz.
- Huss, V. (1978). English word stress in the post-nuclear position. *Phonetica*, 35, 86–105.
- Johnson, K. (2003). *Acoustic and auditory phonetics* (2nd ed.). Malden, MA: Blackwell.

- Kastriani, A. (2003). *The temporal correlates of lexical and phrasal stress in Greek, exploring rhythmic stress: Durational patterns for the case of Greek words*. Master's Thesis, University of Edinburgh.
- Kochanski, G., Grabe, E., Coleman, J., and Rosner, B. (2005). Loudness predicts prominence: Fundamental frequency lends little. *Journal of the Acoustical Society of America*, 118(2), 1038–1054.
- Ladd, R. (1996). *Intonational phonology*. Cambridge: Cambridge University Press.
- Lehiste, I., and Peterson, G. E. (1959). Vowel amplitude and phoneme stress in American English. *Journal of the Acoustical Society of America*, 31, 428–435.
- Lieberman, P. (1960). Some acoustic correlates of word stress in American English. *Journal of the Acoustical Society of America*, 33, 451–454.
- Lindblom, B. (1963). Spectrographic study of vowel reduction. *Journal of the Acoustical Society of America*, 35, 1773–1781.
- Llisterri, J., Machuca, M., de la Mota, C., Riera, M., and Ríos, A. (2003). The perception of lexical stress in Spanish. In M.-J. Solé, D. Recasens, and J. Romero (Eds.), *Proceedings of the XVth International Conference of Phonetic Sciences* (Vol. 2, pp. 2023–2026). Barcelona: Causal Productions.
- Manolescu, A., Olson, D., and Ortega-Llebaria, M. (2009). Cues to contrastive focus in Romanian. In M. Vigário, S. Frota, and M. J. Freitas (Eds.), *Phonetics and phonology: Interactions and interrelations*. Current Issues in Linguistic Theory 306 (pp. 71–90). Amsterdam: John Benjamins.
- Mascaró, J. (2002). Reducció vocàlica. In J. Solà, M.-R. Lloret, J. Mascaró, and M. P. Saldanya (Eds.), *Gramàtica del català contemporani* (pp. 89–123). Barcelona: Editorial Empúries.
- Moon, S.-J., and Lindblom, B. (1994). Interaction between duration, context, and speaking style in English stressed vowels. *Journal of the Acoustical Society of America*, 96(1), 40–55.
- Nakatani, L. H., and Aston, C. H. (1978). Perceiving stress patterns of words in sentences. *Journal of the Acoustical Society of America*, 63 (supplement 1), S55 (A).
- Navarro-Tomás, T. (1914). *Manual de pronunciación española* (3rd ed.). Madrid: Editorial Hernando S. A.
- Navarro-Tomás, T. (1948). *Manual de entonación española* (2nd ed.). New York: Columbia University Hispanic Institute in the United States.
- Navarro-Tomás, T. (1964). La medida de la intensidad. *Boletín del Instituto de Filología de la Universidad de Chile*, 16, 231–235.
- Ortega-Llebaria, M. (2006). Phonetic cues to stress and accent in Spanish. In M. Diaz-Campos (Ed.), *Selected Proceedings of the 2nd Conference on Laboratory Approaches to Spanish Phonology* (pp. 104–118). Somerville, MA: Cascadilla Press.
- Ortega-Llebaria, M., and Prieto, P. (2007). Disentangling stress from accent in Spanish: Production patterns of the stress contrast in de-accented syllables. In P. Prieto, J. Mascaró, and M.-J. Solé (Eds.), *Segmental and prosodic issues in Romance phonology*. Current Issues in Linguistic Theory 282 (pp. 155–176). Amsterdam: John Benjamins.
- Ortega-Llebaria, M., Prieto, P., and Vanrell, M. M. (2008). Perceptual evidence for direct acoustic correlates of stress in Spanish. In J. Trouvain and W. J. Barry (Eds.), *Proceedings of the XVIth International Congress of Phonetic Sciences* (pp. 155–166), Saarbrücken, Germany, 6–10 August 2007.
- Ortega-Llebaria, M., Vanrell, M. M., and Prieto, P. (2010). Catalan speakers' perception of word stress in unaccented contexts. *Journal of the Acoustical Society of America*, 127(1), 462–471.
- Pierrehumbert, J. B. (1980). *The phonetics and phonology of English intonation*. Doctoral Dissertation, MIT.
- Prieto, P., van Santen, J., and Hirschberg, J. (1995). Tonal alignment patterns in Spanish. *Journal of Phonetics*, 23, 429–451.
- Quilis, A. (1971). Caracterización fonética del acento español. *Travaux de Linguistique et de Littérature*, 9, 53–72.
- Quilis, A. (1981). *Fonética acústica de la lengua española*. Biblioteca Románica Hispánica. Madrid: Gredos.
- Quilis, A., and Esgueva, M. (1983). Realización de los fonemas vocálicos españoles en posición fonética normal. In M. Esgueva and M. Cantarero (Eds.), *Estudios de fonética*, I (pp. 159–252). Madrid: Consejo Superior de Investigaciones Científicas.



- Recasens, D. (1986). *Estudis de fonètica experimental del català oriental central*. Barcelona: Publicacions de l'Abadia de Montserrat.
- Recasens, D., and Espinosa, A. (2006). Dispersion and variability of Catalan vowels. *Speech Communication*, 48, 645–666.
- Sluijter, A. M. C. (1995). *Phonetic correlates of stress and accent*. Doctoral Thesis, University of Leiden, Leiden.
- Sluijter, A. M. C., and van Heuven, V. (1996a). Spectral tilt as an acoustic correlate of linguistic stress. *Journal of the Acoustical Society of America*, 100(4), 2471–2485.
- Sluijter, A. M. C., and van Heuven, V. (1996b). Acoustic correlates of linguistic stress and accent in Dutch and American English. *Proceedings of ICSLP 96* (pp. 630–633). Philadelphia, PA: Applied Science and Engineering Laboratories, Alfred I. duPont Institute.
- Sluijter, A. M. C., van Heuven, V., and Pacilly, J. A. (1997). Spectral tilt as a cue in the perception of linguistic stress. *Journal of the Acoustical Society of America*, 101(1), 503–513.
- Solé, M. J. (1984). Experimentos sobre la percepción del acento. *Estudios de Fonética Experimental*, I, 134–243.
- Sorin, C. (1981). Functions, roles and treatments of intensity in speech. *Journal of Phonetics*, 9, 359–374.
- Stevens, K. N. (2000). *Acoustic phonetics*. Cambridge, MA: MIT Press.
- Stevens, S. S. (1971). Perceived level of noise by Mark VII and decibels. *Journal of the Acoustical Society of America*, 51, 575–602.
- Sweet, H. (1906). *A Primer of phonetics*. Oxford: Clarendon.
- Turk, A., and Sawusch, J. (1996). The processing of duration and intensity cues to prominence. *Journal of the Acoustical Society of America*, 99(6), 3782–3790.
- van Heuven, V. (2002). Beyond the segments. Inaugural address at the Royal Netherlands Academy of Arts and Sciences, Amsterdam. Retrieved (17 November 2010) from <http://www.let.leidenuniv.nl/ulcl/faculty/vheuven/oratie-eng.htm>

### Appendix A. Target verb forms in Catalan and Spanish. Underlined syllables receive word stress

	Catalan		Spanish	
	Present 3rd person sing.	Past 3rd person sing.	Present 1st person sing.	Past 3rd person sing.
Verbs	paroxytones	oxytones	paroxytones	oxytones
abominar 'to hate'	ab <u>om</u> ina	abom <u>in</u> à	ab <u>om</u> ino	abom <u>in</u> ó
denominar 'to name'	denom <u>in</u> a	denom <u>in</u> à	denom <u>in</u> o	denom <u>in</u> ó
desanimar 'to discourage'	desan <u>im</u> a	desan <u>im</u> à	desan <u>im</u> o	desan <u>im</u> ó
descaminar 'to cheat'	descam <u>in</u> a	descam <u>in</u> à	descam <u>in</u> o	descam <u>in</u> ó
determinar 'to determine'	determ <u>in</u> a	determ <u>in</u> à	determ <u>in</u> o	determ <u>in</u> ó
discriminar 'to discriminate'	discrim <u>in</u> a	discrim <u>in</u> à	discrim <u>in</u> o	discrim <u>in</u> ó
diseminar 'to spread'	disem <u>in</u> a	disem <u>in</u> à	disem <u>in</u> o	disem <u>in</u> ó
eliminar 'to eliminate'	elim <u>in</u> a	elim <u>in</u> à	elim <u>in</u> o	elim <u>in</u> ó
encaminar 'to guide'	encam <u>in</u> a	encam <u>in</u> à	encam <u>in</u> o	encam <u>in</u> ó
examinar 'to examine'	exam <u>in</u> a	exam <u>in</u> à	exam <u>in</u> o	exam <u>in</u> ó
exterminar 'to kill'	exter <u>min</u> a	exter <u>min</u> à	exter <u>min</u> o	exter <u>min</u> ó
il.luminar 'to light'	il.lum <u>in</u> a	il.lum <u>in</u> à	il.lum <u>in</u> o	il.lum <u>in</u> ó
incriminar 'to incriminate'	incrim <u>in</u> a	incrim <u>in</u> à	incrim <u>in</u> o	incrim <u>in</u> ó
predominar 'to stand out'	predom <u>in</u> a	predom <u>in</u> à	predom <u>in</u> o	predom <u>in</u> ó
recriminar 'to recriminate'	recri <u>min</u> a	recri <u>min</u> à	recri <u>min</u> o	recri <u>min</u> ó

## Appendix B

### Sentences in Catalan

1. Abomina les curses  
*(S)he hates bullfights*
2. No m'agraden gens les curses de braus – abomina la Núria  
*I dislike bullfights – Nuria complains*
3. Abominà la cursa  
*(S)he hated that bullfight*
4. Odio les curses – abominà la Maria  
*I hate bullfights – Maria complained*
5. Determina la massa  
*(S)he establishes its mass*
6. La massa de l'àtom és mesurable – determina complaguda  
*The atom's mass is measurable – she confirms, pleased*
7. Determinà la massa  
*(S)he established its mass*
8. La massa de l'àtom és mesurable – determinà Carmona  
*The atom's mass is measurable – Carmona confirmed*

### Sentences in Spanish

1. Abomino las corridas  
*(S)he hates bullfights*
2. No me gustaron las corridas – abomino contrariada  
*I dislike bullfights – I complain unhappily*
3. Abominó las corridas  
*(S)he hated those bullfights*
4. No me gustaron las corridas – abominó contrariada  
*I disliked bullfights – I complain unhappily*
5. Determino la masa  
*I establish its mass*
6. La masa del átomo es medible – determino complacida  
*The atom's mass is measurable – she confirms, pleased*
7. Determinó la masa  
*(S)he established its mass*
8. La masa del átomo es medible – determinó complacida  
*The atom's mass is measurable – she confirmed, pleased*

## Appendix C

### Desanimar

- a. **Context:** *No sabías si ir a la playa este fin de semana. Catalina te dijo que iba a llover.*  
'You were wondering whether to go to the beach this coming weekend. Catalina told you that it was going to rain.'
- b. **Questions posed by interviewer:**
  - i. *¿Qué pasa? ¿No ibas a la playa?*  
'What's wrong? Weren't you going to the beach?'
  - ii. *¿Qué te dijo Catalina?*  
'What did Catalina tell you?'
- c. **Answers:**
  - i. *Me desanimó Catalina.*  
'Catalina discouraged me.'
  - ii. – *Va a llover – me desanimó Catalina.*  
– 'It's going to rain – Catalina discouraged me.'