Catalan speakers' perception of word stress in unaccented contexts

Marta Ortega-Llebaria
Department of Spanish and Portuguese, University of Texas-Austin (UT), 1 University Station, Austin, Texas 78712

Maria del Mar Vanrell
Department of Translation and Language Sciences, University Pompeu Fabra, C/Roc Boronat, 138, 08018 Barcelona, Spain

Pilar Prieto
Department of Translation and Language Sciences, University Pompeu Fabra, C/Roc Boronat, 138, 08018 Barcelona, Spain and Catalan Institute of Research and Advanced Studies (ICREA), Passeig Lluís Companys, 23, 08010 Barcelona, Spain

(Received 26 October 2008; revised 5 October 2009; accepted 19 October 2009)

In unaccented contexts, formant frequency differences related to vowel reduction constitute a consistent cue to word stress in English, whereas in languages such as Spanish that have no systematic vowel reduction, stress perception is based on duration and intensity cues. This article examines the perception of word stress by speakers of Central Catalan, in which, due to its vowel reduction patterns, words either alternate stressed open vowels with unstressed mid-central vowels as in English or contain no vowel quality cues to stress, as in Spanish. Results show that Catalan listeners perceive stress based mainly on duration cues in both word types. Other cues pattern together with duration to make stress perception more robust. However, no single cue is absolutely necessary and trading effects compensate for a lack of differentiation in one dimension by changes in another dimension. In particular, speakers identify longer mid-central vowels as more stressed than shorter open vowels. These results and those obtained in other stress-accent languages provide cumulative evidence that word stress is perceived independently of pitch accents by relying on a set of cues with trading effects so that no single cue, including formant frequency differences related to vowel reduction, is absolutely necessary for stress perception.
 [DOI: 10.1121/1.3268506]

PACS number(s): 43.71.Es [AJ] Pages: 462–471

I. INTRODUCTION

A historically long line of studies on the acoustic correlates of stress has often led to seemingly contradictory results that prevented a consensus on the common factor that differentiates stressed from unstressed syllables, reflecting the complexity underlying the realization of word stress. Beckman and Edwards (1994) stated that the apparently conflicting results may be because stress cues were examined without controlling for sentence intonation, i.e., pitch accents. A second variable to control is the use that different languages make of vowel reduction in relation to stress (Campbell and Beckman, 1997). Thus, the aim of the present study is to contribute to the body of research on the acoustic realization of word stress by examining the use that speakers of Central Catalan make of duration, overall intensity, and spectral tilt cues to perceive word stress while controlling for the two confounding factors of pitch accents and the formant frequency differences related to vowel reduction.

Following studies of English stress (Bolinger, 1958; Pierrehumbert, 1980; Beckman, 1986; Beckman and Edwards, 1994; Campbell and Beckman, 1997; de Jong, 1991), we assume that stress is used to convey prominence at all levels of the prominence hierarchy, i.e., word stress, phrase stress, and sentence stress, and that this prominence is cumulative across levels. For example, when the sentence "I saw Laura" is spoken as an answer to the question "Who did you see?" (see Fig. 1), the syllable [ˈlɔ] in [ˈlɔɹə] is perceived as stressed because it contains the only open vowel in the word and because, at the sentence level, it bears a pitch accent. Both cues, vowel quality at the word level and pitch accents at the sentence level, reinforce each other to make stress perception more robust.

In stress-accent languages, such as English or Catalan, the presence of a pitch accent is signaled acoustically by the presence of a pitch movement on its landing site, i.e., the stressed syllable, causing co-variation between stress and pitch accents. However, not all syllables with word 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, as Beckman and Edwards (1994) pointed out, this account of English stress may help explain the often contradictory results of studies examining the phonetic correlates of stress since "previous phonetic studies have compared the intensities, durations, and F0 excursions in 'stressed' versus 'unstressed' syllables without controlling systematically for the levels of the stress hierarchy involved" [Beckman and Edwards, 1994, p. 17].

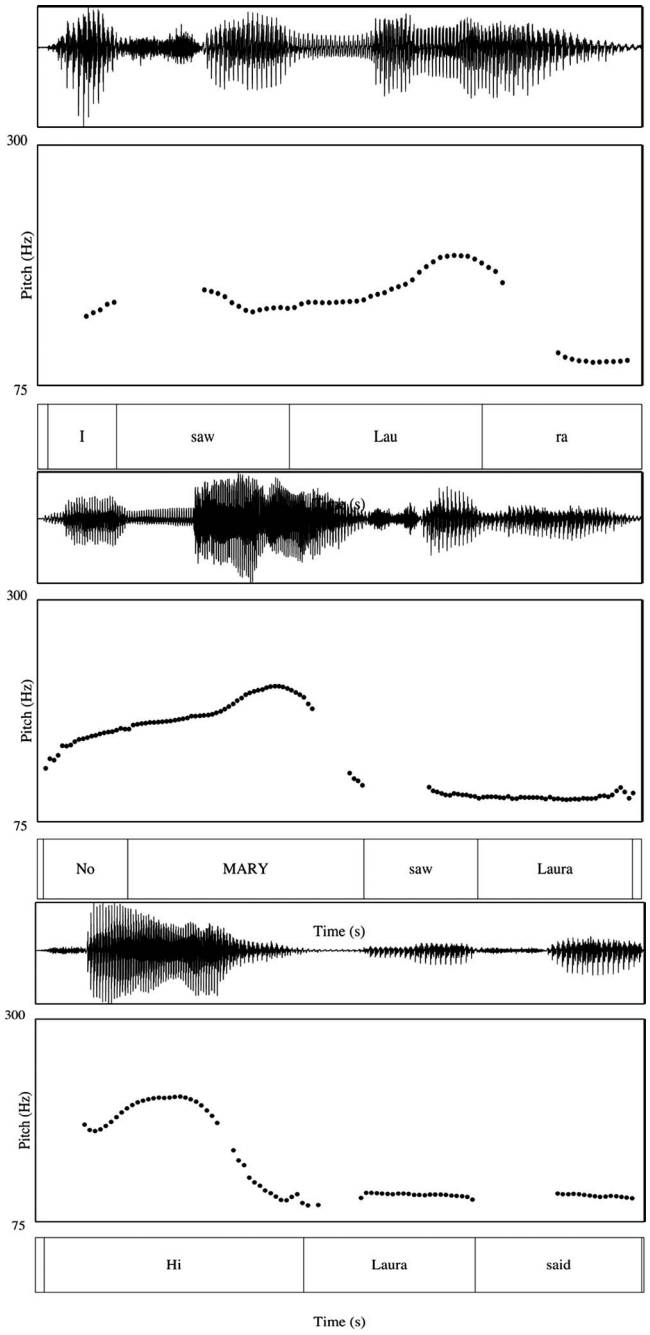

FIG. 1. Declarative sentences, post-focal contexts, and reporting clauses. The word "Laura" aligns with a pitch movement in the declarative sentence (top), while in post-focal contexts (center) and reporting clauses (bottom) "Laura" is produced with a flat *F*0 melody.

Some researchers, however, have controlled for levels of stress by restricting the contrast between stressed and unstressed syllables to the lower levels of the prominence hierarchy; that is, they examined cues to word stress only in unaccented contexts. These were created either by maintaining a constant *F*0 in synthesized stimuli (Fry, 1955, 1958; Turk and Sawusch, 1996), or by asking speakers to produce words in sentences whose intonation does not include pitch accents, such as post-focal contexts (Huss, 1978; Sluijter *et al.*, 1997) or reporting clauses (Ortega-Llebaria *et al.*, 2008). Post-focal contexts, such as "saw Laura" in the sentence "MARY saw Laura" when produced as an answer to "Did Anne see Laura?" (see Fig. 1), are produced with very reduced or no pitch movements, and they are placed just after the focal pitch accent, i.e., "MARY." Reporting clauses are used to report someone's speech, e.g., "He said," "She wrote," and "They shouted," and they are produced with a consistently flat pitch melody and a lower pitch register than the portion of direct speech they introduce (see Fig. 1).

In his pioneering studies on the perception of stress in English, Fry (1955, 1958) found that when listening to noun-verb word pairs that contrasted mainly in stress, i.e., subject noun–subject verb, English speakers relied more on duration cues than on intensity cues to perceive the stress contrast. Turk and Sawusch (1996) complemented Fry's (1955, 1958) results by showing that duration and intensity were perceived integrally, so that irrelevant variations in duration had an effect on the perception of loudness, and irrelevant variations in intensity had an effect on judgments of length. However, this effect was greater for duration, indicating that stress was effectively cued by either a combination of duration and intensity cues, or by duration cues alone, but not by intensity cues only.

The above results contrast with those of Huss (1978), who also examined minimal pairs that differed only in stress, such as import noun–import verb, embedded in post-focal phrases. Although English speakers produced small duration and intensity differences between stressed and unstressed syllables, they identified nouns versus verbs only at chance level. Huss (1978) concluded that in the absence of pitch accents and the formant frequency differences that accompany vowel reduction, duration and intensity cues were insufficient to cue word stress in English. However, these results may be related to the phrasing differences between the target sentences. Although a pause is likely to be inserted after "import" in sentences with a noun target, i.e., [The GERMANS' import] [sinks], a pause is placed before the target word in sentences with the verb target, i.e., [The GERMANS] [import sinks]. This pause causes the last syllable in the phrase to lengthen, making "port" in noun targets to become longer, which reduced the duration differences between nouns and verbs.

Sluijter *et al.* (1997) examined the perception of stress in Dutch by embedding [ˈnana] and [naˈna] in post-focal contexts, while controlling for phrasing effects. Similar to the English speakers in Fry's (1958) and Turk and Sawusch's (1996) experiments, Dutch speakers perceived stress by relying on duration and intensity cues. However, intensity differences related to word stress were perceived better with changes in spectral tilt, i.e., intensity differences between the higher and lower regions of the spectrum, than with changes in overall intensity, i.e., the mean intensity across all frequencies of the spectrum.

Ortega-Llebaria *et al.* (2008) embedded the words *mama* and *mimi* in reporting clauses to examine the perception of stress in Castilian Spanish, a language that, in contrast with Dutch and English, has no phonological vowel reduction. Spanish speakers perceived the stress differences between [ˈmama]-[maˈma] and [ˈmimi]-[miˈmi] by relying mainly on duration cues, and to a lesser extent, in overall intensity cues. Speakers did not use spectral tilt cues.

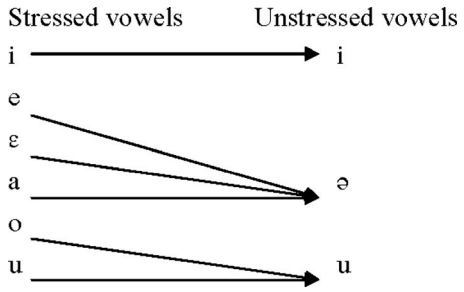

FIG. 2. Vowel inventory of Central Catalan. Lines indicate patterns of vowel reduction.

In general, the above studies provide cumulative evidence that word stress is perceived in the absence of pitch accents across languages by relying on duration. Stressed syllables are consistently longer than their unstressed counterparts. In contrast to duration, languages differ in their use of intensity cues; i.e., Spanish speakers use overall intensity while Dutch speakers rely on spectral tilt. However, these experiments limited their test materials to bi-syllabic words with open vowels. Although this constitutes a realistic strategy to stress perception in languages that have no vowel reduction, such as Spanish, in languages such as English or Dutch, where words alternate stressed open vowels with unstressed mid-central vowels, duration and intensity cues to word stress should also be examined in relation to vowel quality.

Campbell and Beckman (1997) highlighted the important role of vowel reduction in stress perception by suggesting that the different use that languages make of duration and spectral tilt cues to mark the stress contrast is a direct consequence of their differences in the degree of co-variation between stress and the formant frequency differences linked to vowel reduction. More specifically, they could not replicate Sluijter *et al.*'s (1997) results for Dutch in English, since unlike in Dutch, duration in English was not a consistent cue to stress and variations in spectral tilt correlated with pitch accents rather than stress proper. They argued that the different results obtained for Dutch and English are related "with differences between the two languages in the perceptibility of stress in the absence of accent or vowel reduction […] Dutch differs from English in having relatively fewer words in which unstressed syllables are reduced, particularly in word initial position" [Sluijter *et al.*, 1997, p. 70].

The formant frequency differences related to vowel reduction constitute a reliable cue to stress only in English, and, therefore, English speakers, unlike Dutch speakers, do not need to rely on other cues, such as duration or spectral tilt, to perceive the stress contrast. However, because these differences in stress and vowel reduction take place across languages, the different use that Dutch and English speakers make of duration and spectral tilt cues to stress could be related to other cross-language differences besides vowel reduction. Vowel reduction patterns in Central Catalan provide an appropriate context to examine different degrees of co-variation between stress and formant frequency differences between corresponding vowels within a single language. As shown in Fig. 2, Central Catalan allows seven vowels [i, e, ɛ, a, ɔ, o, u] to appear in stressed position and three vowels [i, ə, u] to appear in unstressed position (Wheeler, 2005; Herrick, 2003; Recasens, 1986; Recasens and Espinosa, 2006). In an unstressed position, the vowels /i/ and /u/ surface as [i] and [u], respectively (*vi* [ˈbi] "wine" > *vinet* [biˈnɛt] "wine.dim;" *suc* [ˈsuk] "juice" > *suquet* [suˈkɛt] "juice.dim") (Recasens, 1986, p. 131), the vowels /e, ɛ, a/ merge and surface as schwa [ə] (*peix* [ˈpeš] "fish" > *peixet* [pəˈšɛt] "fish.dim;" *sec* [ˈsek] "dry" > *sequet* [səˈkɛt] "dry.dim;" *sac* [ˈsak] "bag" > *saquet* [səˈkɛt] "bag.dim"), and the vowels /ɔ, o/ merge and surface as [u] (*poc* [ˈpɔk] "few" > *poquet* [puˈkɛt] "few.dim;" *boca* [ˈbokə] "mouth" > *boqueta* [buˈkɛtə] "mouth.dim"). Consequently, it is possible to have words in which vowel quality co-varies with stress, such as [ˈbokə] "mouth," [pəˈšɛt] "fish dim.," and [bəˈlentə] "courageous fem.," as well as words that contain no vowel quality cues to stress, i.e., [biˈžili] "watch out," [ˈlikit] "liquid," and [impulˈsiu] "impulsive."

Duration, overall intensity, and spectral correlates of word stress were measured in Catalan-Spanish cognates embedded in accented and in unaccented sentences in a pilot production experiment. Cognates consisted of four-syllable verbs that ended in [ˈminə] and [miˈna] in Catalan and in [ˈmino] and [miˈno] in Spanish. Measurements of target *–mina* and *–mino* endings showed that formant frequency differences between corresponding vowels took place only in Catalan [a], where *F*1 was 200 Hz higher than in [ə], its unstressed counterpart. No vowel quality differences in relation to stress or accent were found for Spanish [o] and [i], and for Catalan [i]. Moreover, no significant vowel quality differences were found between Catalan [i] and Spanish [i]. Speakers of both languages produced stressed vowels with durations 11 ms longer and overall intensities 1.7 dB greater than their unstressed counterparts in both accented and unaccented contexts. However, duration differences were significantly greater in Catalan [a] than in other vowels, suggesting that the [a]-[ə] alternation amplified the duration differences related to the stress contrast. No significant differences in spectral tilt were found in Catalan [i] and Spanish [o] and [i], the vowels that maintained the same quality across stress contexts. However, a gain of 4 dB differentiated the spectral tilt in stressed [a] from that in unstressed [ə], indicating that in our production data, spectral tilt correlated with formant frequency differences related to stress rather than stress proper. Spanish speakers' perception of word stress in unaccented contexts (Ortega-Llebaria *et al.*, 2008) confirmed that Spanish speakers perceived the stress contrast in [a] and [i] in the absence of pitch accents and vowel reduction patterns by relying on duration and overall intensity cues, but not on spectral tilt.

The goal of the present experiment is to examine how speakers of Central Catalan use duration, overall intensity, and spectral tilt cues to perceive word stress in unaccented contexts. In order to control for possible effects on stress of the formant frequency differences related to vowel reduction, these cues are examined in words with no vowel quality cues to stress and in words that contained open and mid-central vowels. Given results reported in the production literature, it

is of particular interest to examine how duration interacts with vowel reduction, and whether duration and spectral tilt differences to stress are perceived only in words that alternate open and mid-central vowels or also in words with no vowel quality cues to stress.

II. METHODOLOGY

A. Experimental design and statistics

In order to examine which cues and cue combinations Catalan speakers use to perceive lexical stress in unaccented contexts, Catalan subjects were asked to listen and identify the word *mamà* (or *mimí*), with stress on the last syllable, in a reporting clause. Reporting clauses are a particular clause type used to report someone's speech, (e.g., "He said," "She wrote," and "They shouted"), and they are produced with a flat pitch melody and a lower pitch register than the direct speech they introduce. They do not present any tonal movement, and in sentences up to ten syllables, there are no internal divisions or phrasing effects (for a detailed description of *F*0 in reporting sentences in Catalan, see Astruch, 2005). A native speaker of Central Catalan produced these items (Sec. II B), which were manipulated to create stimuli that combined various values of duration for the two syllables of *mama* (and *mimi*) in order to obtain a continuum that ranged from ratios typical of words with penultimate stress, i.e., [ˈmama] or [ˈmimi], to ratios of words with last syllable stress, i.e., [maˈma] or [miˈmi]. Similar continua were obtained for values of overall intensity and spectral tilt in *mama* and *mimi* words. Since in Central Catalan, [i] maintains similar vowel qualities across different stress and accent contexts while stressed open [a] alternates with an unstressed mid-central [ə] (Recasens, 1986; Recasens and Espinosa, 2006; Herrick, 2003), only *mama* was manipulated for vowel reduction (Sec. II C). Values of the duration continuum were crossed with those of overall intensity, spectral tilt, and the vowel reduction continua yielding grids typical of an expanded factorial design, i.e., the "duration × overall intensity" grid consisted of the 25 stimuli obtained by crossing the 5 levels of the duration continuum with the 5 levels of the overall intensity continuum (see, for instance, Massaro *et al.*, 1993; Chen and Massaro, 2004). Stimuli from each grid were presented in an identification task. The proportions of *mamà* (or *mimí*) identification answers were subjected to logistical regression analyses (GENMOD model in SAS statistics package) in order to obtain GEE parameter estimates and odd ratios for each cue to stress within each identification task. Significant interactions were further analyzed by calculating the regression equation within each level of duration, spectral tilt, and vowel reduction.

B. Recordings

A 41 year-old female native speaker of Central Catalan was recorded saying the sentence *Hola – saluda la mama contenta* "Hi – greets mom happily," where the reporting clause *saluda la mama contenta* was consistently pronounced with a flat pitch melody. The target word [ˈmamə] was replaced with [məˈma], [miˈmi], and [ˈmimi], yielding a total of 60 sentences (4 target words × 15 repetitions).

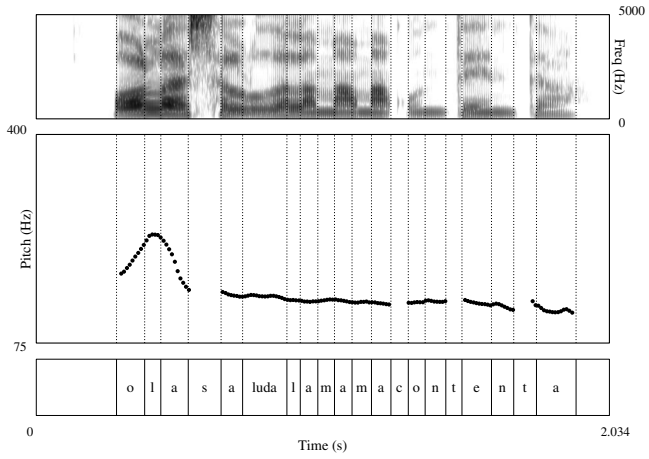

FIG. 3. Spectrogram, pitch track, and labels of the sentence *Hola – Saluda la mama contenta* "Hi – greets mom happily." Notice the pitch accent on "Hola" and the flat *F*0 melody for the reporting clause.

Measurements of duration, intensity, and spectral tilt were made on all target words, and the sentences containing the [ˈmamə] and [ˈmimi] items with values closest to the average were selected for further manipulation.

C. Materials

1. Duration and intensity

In order to examine the role of duration and intensity in the perception of primary stress, we controlled for formant frequency differences linked to vowel reduction. Recall that speakers from Central Catalan alternate stressed open [a] with unstressed mid-central [ə]. Thus, in order to avoid vowel reduction in the target word [ˈmamə], the first stressed vowel was copied to the second syllable, yielding two vowels with identical spectral characteristics, i.e., [mama]. For consistency, the same manipulations were performed on *mimi*. Afterwards, the target words were inserted in the carrier sentence, and, when necessary, their pitch was manipulated with the PSOLA utilities in PRAAT (Boersma and Weenink, 2005) to ensure that they had a nearly flat pitch trajectory. *F*0 values at the beginning, midpoint, and end of the target vowels were 138.5, 137.6, and 136.01 Hz for the first vowel of *mama*; 137.8, 138.04, and 137.9 Hz for the second vowel; and 135.5, 135.6, 135.5 Hz and 134, 134.4, and 133.5 Hz for the first and second vowels of *mimi* (Fig. 3).

A total of three five step [ˈmama]-[maˈma] continua (and three [ˈmimi]-[miˈmi] continua) were created by separately manipulating the cues of duration, overall intensity, and spectral tilt. For each continuum, stimulus 1 had a syllable ratio typical of words with penultimate stress, i.e., [ˈmama]. This ratio decreased in stimulus 2, was close to 1 in stimulus 3, and increased again but in the opposite direction for stimuli 4 and 5, with stimulus 5 replicating the ratio of words with last syllable stress, i.e., [maˈma]. For example, in the overall intensity continuum, syllable 1 was 3 dB louder than syllable 2 in stimulus 1. This difference decreased to 1.5 dB in stimulus 2, and became close to 0 in stimulus 3. In stimulus 4 the second syllable was 1.5 dB louder than the first, and this difference increased to 3 dB in

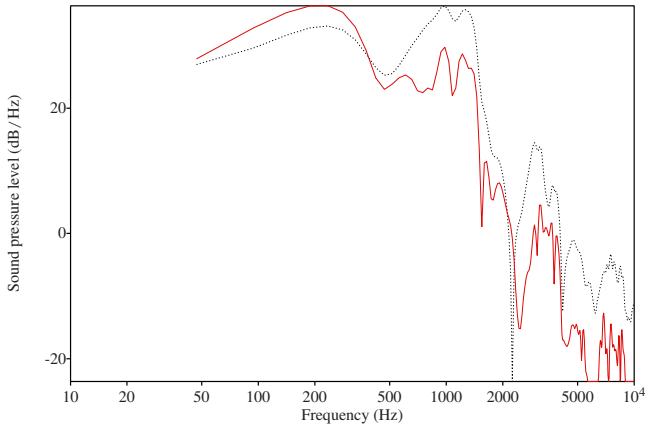

FIG. 4. (Color online) Spectral tilt of the vowels in *mama*. In stimulus 1 on the spectral tilt continuum, frequencies above 500 Hz have been amplified by 8 dB in the first vowel of *mama* producing a flatter spectral tilt in this vowel (dotted line) when compared to the spectral tilt of second vowel (compact line).

stimulus 5. These increments, which were achieved by multiplying all frequencies of the selected syllable by a constant factor with the multiply function in PRAAT (Boersma and Weenink, 2005), were modeled after our pilot production experiment, where Catalan stressed syllables in unaccented contexts were an average 1.57 dB (SD=1.81) louder than their unstressed counterparts.

In the spectral tilt continuum, amplitude differences between the lower and higher regions of the spectrum were achieved by manipulating the amplitude of the frequencies above 500 Hz in one of the *mama* (or *mimi*) vowels. Amplitude was increased to 4 dB by multiplying all frequencies above the 500 Hz range by a common factor using the finite impulse response filters with frequency bands spaced one-third of octave in the ADOBE AUDITION software. Then, the whole vowel was leveled for intensity with the scale intensity function in PRAAT (Boersma and Weenink, 2005) in order to maintain the same overall intensity as the other vowel in *mama* (or *mimi*). In stimulus 1, amplitude differences between frequencies below and above 500 Hz in the first vowel of *mama* (or *mimi*) were 8 dB greater than those in the second vowel (see Fig. 4). This difference was 4 dB in stimulus 2. Both vowels in stimulus 3 had identical spectral tilts, and in stimuli 4 and 5 the amplitude differences below and above 500 Hz were 4 and 8 dB greater in the second vowel than in the first vowel of *mama* (or *mimi*).

Boosting the amplitude of frequencies over 500 Hz was based on Sluijter and van Heuven's (1996) and Sluijter *et al.*'s (1997) experiments for Dutch. They found that intensity level differences between stressed and unstressed Dutch syllables were concentrated in the higher parts of the spectrum, whereas intensity differences below 500 Hz were negligible. In our stimuli, because formant values for [a] were above 500 Hz (*F*1: 869 Hz, *F*2: 1315 Hz, and *F*3: 3102 Hz) while for [i], *F*1 was below 500 Hz (*F*1: 324 Hz, *F*2: 2686 Hz, and *F*3: 3130 Hz), these manipulations boosted the amplitudes of *F*1, *F*2, and *F*3 with respect to *F*0 in [a], and *F*2 and *F*3 with respect to *F*0 and *F*1 in [i].

The value of the amplitude increments was fixed at 4 dB after measuring the Catalan speakers' productions in our pilot production experiment. On average, amplitude differences between frequencies below and above 500 Hz were 2.87 dB (SD=5.8) in stressed [a] and 7.01 dB (SD=5.65) in unstressed [ə], yielding an approximate gain of 4.14 dB for stressed [a] over its unstressed counterpart. In contrast, speakers produced remarkably similar spectral tilts for stressed and unstressed vowel [i] (M=6.56 dB and SD =5.73 for stressed [i], and M=6.79 dB and SD=5.51 for unstressed [i]).

Duration ratios between the two syllables of [mama] (or [mimi]) were modified by manipulating vowel duration while maintaining word duration of [mama] constant at 267 ms (or at 259 ms in [mimi]). Starting from stimulus 1, durations in milliseconds for each vowel in [mama] were 84-56, 77-63, 70-70, 63-77, and 56-84, respectively. In [mimi], durations in milliseconds were 77-49, 70-56, 63-63, 56-70, and 49-77. Durations were modified by cutting glottal cycles from one of the vowels in [mama] (or [mimi]) and adding cycles to the other vowel. Duration values were based on those obtained in our pilot production experiment. Stressed [a] was an average 82.5 ms long (SD=13.01) while unstressed [ə] was 61.9 ms (SD=10.15) yielding an average difference of 20.41 ms (SD=8.96). Stressed [i] was an average 58.56 ms long (SD=9.63) and unstressed [i] was 50.16 ms (SD=10.73) yielding a 6.39 ms difference (SD=8.26).

Finally, the five levels of the duration continuum were crossed with those of the overall intensity continuum, creating a 5×5 grid for [mama] and [mimi]. For example, when crossing overall intensity 1 with duration 1 of the continua, the resulting [mama] stimulus had a first vowel 84 ms long, a second vowel 56 ms long, and the first vowel was 3 dB higher than the second vowel. Similarly, the five levels of

TABLE I. Summary of identification tasks. There were five identification tasks, one per grid. Each grid contained the *mama* (or *mimi*) stimuli resulting from crossing different continua.

Crossed continua	Stimuli	No. of repetitions per stimulus	Subjects
Five step duration×five step overall intensity	25 mama	7	Group A=10 subjects
	25 mimi	7	Group B=10 subjects
Five step duration×five step spectral tilt	25 mama	7	Group A=10 subjects
	25 mimi	7	Group B=10 subjects
Seven step duration×two step vowel quality	14 mama	10	Group A+group B=20 subjects

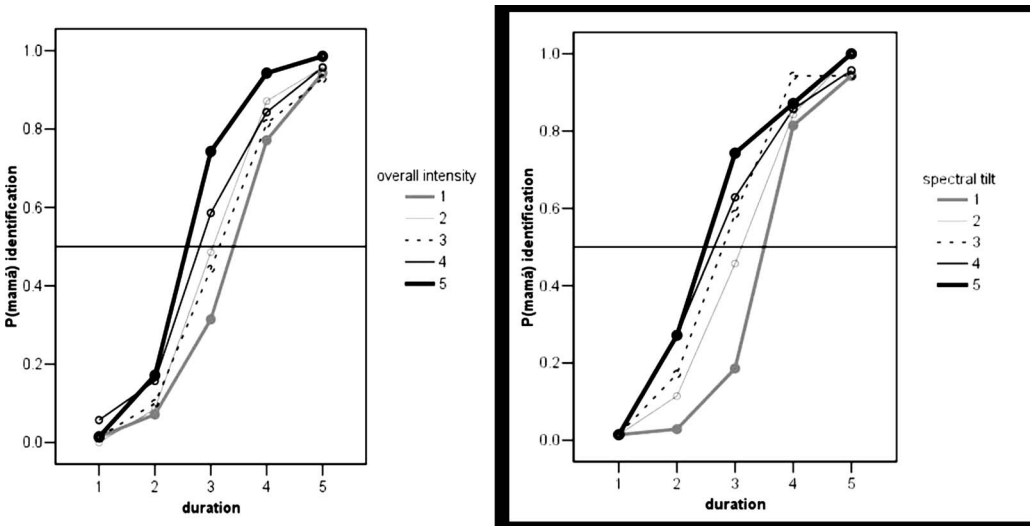

FIG. 5. Proportion of *mamà* responses for the 25 stimuli resulting from crossing the 5 levels of the duration continuum with the 5 levels of the overall intensity continuum in (a) [or the 5 levels of duration continuum with the 5 levels of spectral tilt continuum in (b)]. Duration and intensity ratios [or duration and spectral tilt ratios in (b)] reflect paroxytone stress in stimulus 1, i.e., [ˈmama], and oxytone stress in stimulus 5, i.e., [maˈma].

duration continuum were also crossed with the five of spectral tilt. A summary of the crossed continua is depicted in Table I.

2. Duration and vowel quality

In order to examine the effect of duration and vowel quality on the perception of primary stress, one [ˈməma] token and one [məˈma] token were selected for further manipulation. We made sure that the formant values for the stressed vowel in [ˈməma] showed an open quality (*F*1: 869 Hz, *F*2: 1315 Hz, and *F*3: 3102 Hz) while the values for the unstressed vowel in the second syllable were typical of the mid-central schwa (*F*1: 657 Hz, *F*2: 1237 Hz, and *F*3: 2944 Hz). In [məˈma], the vowel quality patterns reversed, yielding schwa values for the first vowel (*F*1: 668 Hz, *F*2: 1184 Hz, and *F*3: 2981 Hz) and an open quality for the second vowel (*F*1: 832 Hz, *F*2: 1486 Hz, and *F*3: 3040 Hz).

The vowel duration of each of these two tokens was manipulated to create two seven step continua while duration of [ˈməma] and [məˈma] remained constant at 235 ms. These manipulations aimed to have stimuli with short open [a] and long mid-central [ə] in addition to the stimuli with long open [a] and short mid-central [ə]. Starting from stimulus 1, the durations in milliseconds of the first and second vowels in [ˈməma] were 87-45, 80-52, 73-59, 66-66, 59-73, 52-80, and 45-87. A second continuum was created based on [məˈma]. Durations in milliseconds of the first and second vowels, starting from stimulus 1, were 45-85, 52-78, 59-71, 66-64, 73-57, 80-50, and 87-43.

In contrast to [a], Catalan [i] maintains the same vowel quality across stress contexts (Recasens, 1986, p. 131), which precludes Catalan listeners from identifying vowel nuclei with the quality of a schwa as possible realizations of Catalan [i]. Therefore, a duration × vowel quality identification task with [mimi] stimuli was not performed.

D. Subjects and identification tasks

20 native speakers of Central Catalan participated in the study. Their ages ranged from 19 to 41 years old, and they were born in areas where Central Catalan was spoken or had lived there for most of their lives. Catalan was the language they spoke with their families and friends, and they considered it their mother tongue. They had learned Spanish as a second language later in school, starting from age 6. None of them reported having any speech or hearing problems.

A group of ten subjects (group A in Table I) was instructed that they would hear the words *mamà* "mother," which is spoken with stress on the last syllable, or *mama* "mommy," which is pronounced with penultimate stress. Catalan orthographic conventions distinguish these two words by placing a stress mark in the stressed vowel of [məˈma], e.g., *mamà*. Subjects were asked to press the space bar in a keyboard as soon as they heard the word *mamà* in the sentence *Hola – saluda la ___ contenta* "Hi –greets _____ happily" over headphones. They listened to the randomized 175 sentences of the "duration × overall intensity" condition (25 stimuli of grid 1 × 7 repetitions) in 7 blocks of 25 stimuli with an inter stimuli interval (ISI) of 1500 ms and a 10-s pause between blocks. After a longer rest, they then listened to the 175 sentences of the "duration × spectral tilt" condition (25 stimuli of grid 3 × 7 repetitions). Presentation orders between the two tasks were counterbalanced across subjects. The remaining ten speakers (group B in Table I) received the same instructions that subjects in group A performed analogous tasks with the [ˈmimi]-[miˈmi] stimuli and spelled *mimi* and *mimí* in Catalan orthography. Finally, all 20 subjects (groups A and B) listened to the 10 repetitions of the 14 "duration × vowel quality" stimuli.

III. RESULTS

A. Duration, overall intensity, and spectral tilt in *mama*

The graphs in Fig. 5 illustrate the proportions of *mamà* answers for the 25 stimuli resulting from crossing the 5 steps of the duration with the 5 steps of the overall intensity continua [Fig. 5(a)], or with the 5 steps of the spectral tilt continuum [Fig. 5(b)]. Both graphs are similar in that the slopes

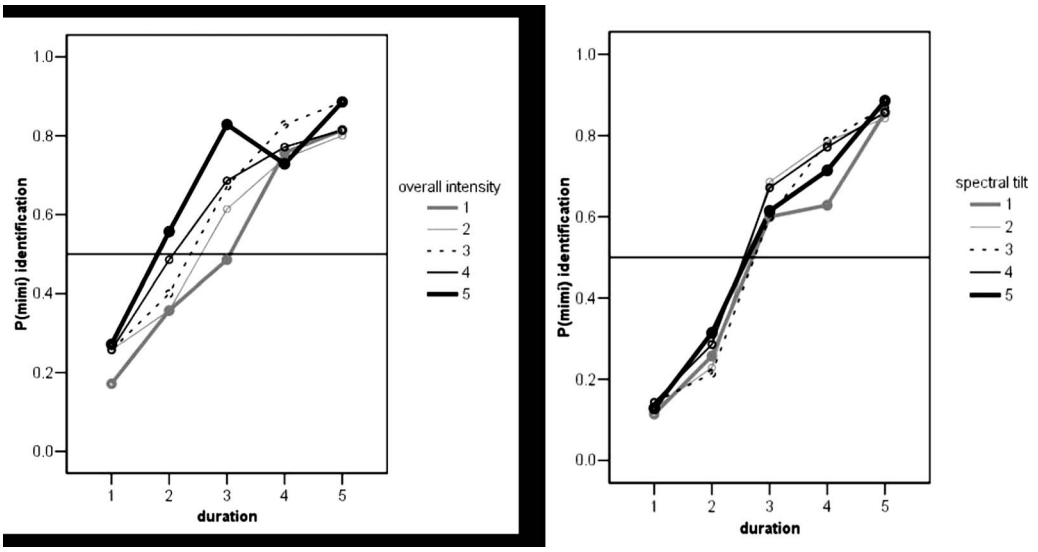

FIG. 6. Proportions of *mimí* responses for the 25 stimuli resulting from crossing the 5 levels of the duration continuum with the 5 levels of the overall intensity continuum in (a) [or the 5 levels of duration continuum with the 5 levels of spectral tilt continuum in (b)]. Duration and intensity ratios [or duration and spectral tilt ratios in (b)] reflect paroxytone stress in stimulus 1, i.e., [′mimi], and oxytone stress in stimulus 5, i.e., [mi′mi].

of the functions indicate that duration has a strong effect on the perception of primary stress: For each intensity level [measured as overall intensity in Fig. 5(a) or as spectral tilt changes in Fig. 5(b)], the [ma′ma] responses increase along the duration continuum regardless of whether they are in a competing or enhancing relationship with intensity. A second similarity between these two graphs is that the spread among curves, especially in duration stimuli 2 and 3, indicates that intensity also has an effect on the perception of stress.

Logistic regression analyses corroborate these visual trends. Analysis of GEE parameter estimates show a significant effect of duration ($Z=-8.97$, $p<0.0001$) and overall intensity ($Z=-2.36$, $p=0.018$) in Fig. 5(a), and of duration ($Z=-11.83$, $p<0.0001$), spectral tilt ($Z=-4.33$, $p<0.0001$), and duration × spectral tilt ($Z=2.01$, $p=0.044$) in Fig. 5(b), confirming that Catalan speakers use all the cues present in the stimuli to make stress judgments. Moreover, odd ratios of parameter estimates give us a reasonable estimate of cues' strength because the duration and intensity values used in this perception experiment approximate those obtained in our pilot production experiment, which averaged measurements on 1200 vowels produced by 20 speakers. They indicate that listeners rely more strongly on duration than on intensity cues. As duration increases by one whole unit in Fig. 5(a), listeners are 6.17 times more likely to perceive *mamà*, while increments in overall intensity show only a 1.34 increase in likelihood. Likewise, at each duration increment in Fig. 5(b), listeners are 8.39 times more likely to perceive *mamà*, but spectral tilt increments yield only a 1.9 likelihood increase.

The significant interaction between duration and spectral tilt was further analyzed. By investigating spectral tilt at each level of duration, it was found that spectral tilt has a significant effect only for durations 2 and 3 ($Z=-4.34$ and $Z=-6.67$ at $p<0.0001$; non-significant effect for durations 1, 4, and 5 scored $Z=-0.09$, $Z=-1.07$, and $Z=-2.1$ at $p>0.05$), in which both syllables of *mama* have similar or identical durations, showing that listeners rely on spectral tilt cues especially in those stimuli in which duration cues to stress are ambiguous. In contrast, when duration is examined at each level of spectral tilt, duration cues have a significant effect across the five levels of spectral tilt (starting from level 1, $Z=-8.03$, $Z=-15.69$, $Z=-7.76$, $Z=-13.42$, and $Z=-9.00$, with $p<0.0001$). Thus, Catalan speakers consistently rely on duration cues to perceive stress, not only for stimuli where spectral cues to stress become ambiguous, but also in those where spectral tilt is in either an enhancing or a competing relationship with duration.

The wider spread of the overall intensity curves for duration 3 in Fig. 5(a) is also examined by exploring the effect of overall intensity at each level of duration. Overall intensity has a significant effect only for duration 3 ($Z=-4.19$, $p<0.0001$; non-significant results for durations 1, 2, 4, and 5: $Z=-1.26$, $Z=-3.3$, $Z=-2.39$, and $Z=-1.46$ at $p>0.05$), indicating that listeners' use of overall intensity cues becomes evident when duration cues are ambiguous in the stimuli.

In summary, listeners rely heavily on duration to identify word stress, but the addition of overall intensity or spectral tilt as enhancing cues results in a shift of the response profile toward the alternative *mamà*, showing an additive relationship of duration with overall intensity and spectral tilt cues. However, listeners still perceived stress consistently in stimuli that contained duration cues alone, and also in those that contained spectral tilt cues alone, or overall intensity alone.

B. Duration, overall intensity, and spectral tilt in *mimi*

The graphs in Fig. 6 show the proportions of *mimí* answers for the 25 stimuli resulting from crossing the 5 steps of the duration with the 5 steps of the overall intensity continua [Fig. 6(a)], or with the 5 steps of the spectral tilt continua [Fig. 6(b)]. As in the graphs for *mama*, these graphs for *mimi* show that *mimí* responses increase along the duration continuum at each overall intensity [Fig. 6(a)] and spectral tilt

[Fig. 6(b)] curve, showing a consistent effect of duration cues in the perception of stress. However, unlike in *mama*, judgments at the end points are not unanimous possibly because of differences between subjects' groups. The spread of the curves is larger in Fig. 6(a) than in Fig. 6(b), indicating that overall intensity cues may have a stronger effect on the perception of stress in *mimi* than spectral tilt cues.

Analysis of GEE parameter estimates confirm the significant effect of duration ($Z=-4.60$, $p<0.0001$) and overall intensity ($Z=-2.93$, $p=0.003$) in Fig. 6(a), and of duration ($Z=-5.30$, $p<0.0001$) in Fig. 6(a). However, spectral tilt, as well as the interactions of duration with overall intensity and of duration with spectral tilt, yields non-significant results (spectral tilt: $Z=-0.29$, $p=0.77$; duration × intensity: $Z=1.17$, $p=0.24$; and duration × spectral tilt: $Z=-0.15$, $p=0.97$). Thus, Catalan listeners use duration and overall intensity cues to perceive lexical stress in *mimi*, but no significant effect is found for spectral tilt cues. Moreover, listeners rely more on duration than on overall intensity cues, because odd ratios of parameter estimates show that the probability of hearing *mimí* increases 2.17 times along each duration step, but it increases only 1.23 times at each overall intensity step. Finally, logistic regression analyses run separately at each duration level showing that overall intensity had a significant effect only for durations 2 and 3 ($Z=-3.44$, $p=0.0006$ and $Z=-3.76$, $p=0.0002$, respectively; non-significant effects at durations 1, 4, and 5: $Z=-1.53$, $Z=0.24$, and $Z=-2.28$ at $p<0.05$), while spectral tilt had non-significant effects at all levels at $p>0.05$ ($Z=-0.45$, $Z=-0.88$, $Z=-0.06$, $Z=-1.42$, and $Z=-0.58$).

When listening to the stress contrast in *mimi* stimuli, Catalan speakers relied strongly on duration, and to a lesser extent, overall intensity. Listeners still perceived the stress contrast in stimuli that contained duration cues alone or overall intensity cues alone. Spectral tilt had no significant effect.

C. Duration and vowel reduction in *mama*

Figure 7 shows the proportions of *mamá* answers for the 14 stimuli resulting from crossing the 7 step duration continuum with the [mamə] and [məma] items. Listeners identified *mamá* more often in the [məma] item (black line) than in the [mamə] item (gray line), showing that, in general, vowels with open qualities, i.e., [a], are perceived as stressed and vowels with mid-central qualities, i.e., [ə], are perceived as unstressed. However, the slopes of the curves showed that duration also had an effect on the perception of stress. Listeners perceived more *mamà* items when the second vowel was longer (stimuli 5–7) even in [mamə], where the longer vowel has a mid-central quality. In fact, listeners identified 58% of [mamə] items as *mamà* on duration 5, 68% on duration 6, and 66% on duration 7 while only 0.7% of [mamə] items were heard as *mamà* when [a] was longer than [ə]. They also heard fewer *mamà* items in stimuli 1–3 even in the [məma] continuum, where again [ə] is longer than [a]. For example, only 37% of [məma] items were identified as *mamà* in duration 1 against 97% in duration 7.

An analysis of GEE parameter estimates indicates that duration, vowel quality, and their interaction had a significant

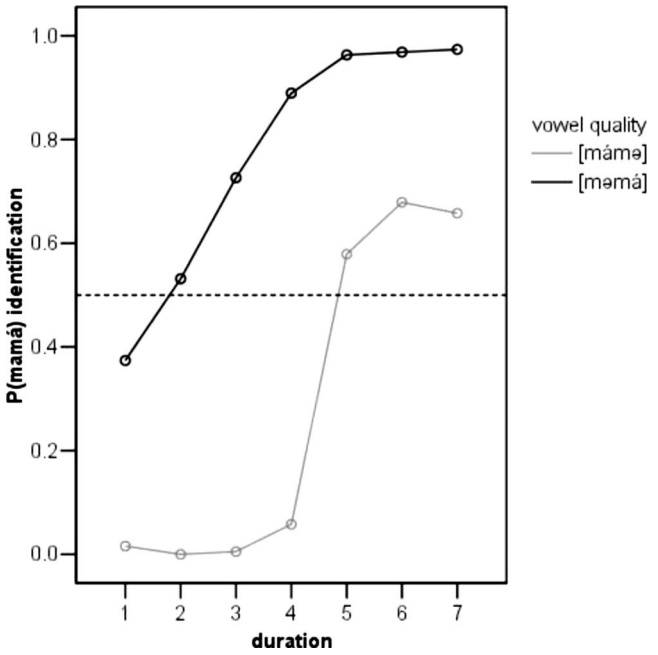

FIG. 7. Proportions of *mamà* responses for the 14 stimuli resulting from crossing the 7 levels of duration with the 2 levels of vowel quality in a [′mamə]-[mə′ma] continuum. Starting from stimulus 1, the durations in milliseconds of the first and second vowels in [′mamə] were 87-45, 80-52, 73-59, 66-66, 59-73, 52-80, and 45-87; and in [mə′ma] were 45-85, 52-78, 59-71, 66-64, 73-57, 80-50, and 87-43.

effect on the perception of stress ($Z=-11.48$, $p<0.0001$ for duration; $Z=-9.55$, $p<0.0001$ for vowel quality; and $Z=2.14$, $p=0.032$ for their interaction). The interaction was further analyzed by applying the logistic regression separately at each duration level and at each level of vowel reduction. Results showed a significant effect of vowel quality in each level of duration at $p<0.01$ ($Z=-5.83$, $Z=-5.97$, $Z=-6.12$, $Z=-8.95$, $Z=-7.68$, $Z=-7.78$, and $Z=-6.39$), and crucially, a significant effect of duration in the two levels of vowel quality ($Z=-9.43$ and $Z=-11.48$ at $p<0.01$), confirming, on the one hand, that the alternation of open and mid-central vowels had a strong effect on listeners' stress judgments. On the other hand, duration modified the perception of stress in [mamə] and [məma] items so that longer mid-central vowels were perceived as more stressed than shorter open vowels.

In summary, listeners consistently relied in both vowel quality and duration to perceive stress. Longer vowels with open qualities were perceived as more stressed than shorter vowels with mid-central qualities. However, in stimuli with conflicting cues, longer mid-central vowels were frequently perceived as more stressed than short open vowels, showing that the formant frequency differences related to vowel reduction did not inhibit duration cues to stress prominence.

IV. DISCUSSION

Speakers of Central Catalan perceived word stress in unaccented contexts in words that had no vowel quality cues to stress. More specifically, speakers used duration and overall intensity cues as effectively in strings of open vowels that were common in Catalan (i.e., [mimi]) as in those combinations of open vowels that were not possible Catalan words

(i.e., *[mama] instead of [mamə] or [məma]). Similar results have been obtained in languages such as English (Fry, 1955, 1958; Turk and Sawusch, 1996), Dutch (Sluijter *et al.*, 1997), and Spanish (Ortega-Llebaria *et al.*, 2008), showing that word stress is perceived cross-linguistically in the absence of pitch accents and vowel quality cues to stress by relying on duration and overall intensity cues. Thus, duration and intensity cues are used by speakers of languages that differ greatly in their degree of co-variation between vowel reduction and stress, that is, languages such as English, with a strong co-variation between stress and vowel quality as well as languages such as Spanish, where vowels maintain the same quality across stress contexts or Catalan, where only some stressed open vowels alternate with unstressed mid-central vowels.

Moreover, the present results show that Catalan speakers relied strongly on duration to perceive stress in words that alternated vowels with open and mid-central qualities as well. When the duration continuum was crossed with [mamə] and [məma] stimuli, Catalan speakers relied on both vowel quality and duration cues to perceive stress, showing that these cues were in an additive relationship. In stimuli with conflicting cues, listeners perceived longer [ə] as more stressed than shorter [a] in both [mamə] and [məma] items, revealing that not only vowels with open qualities are perceived as stressed, but that mid-central vowels of sufficient length also induce the perception of stress. Similar perception patterns have been found in European Portuguese, where Delgado Martins (1975) showed that duration affects stress perception not only in words such as *splicitu*, where there are no vowel quality cues to stress, but also in words such as *splicito* when the full [o] is shortened more than [i]. Interestingly, Fry (1964) showed a similar effect on English. He crossed a duration continuum with a vowel reduction continuum and found that English speakers relied more strongly on duration than on vowel quality cues to perceive word stress. He cautioned that his results were related to the different scales used in the two continua. Thus, if vowels with non-peripheral, mid-central qualities are *sufficiently* long, they induce stress perception not only in languages such as Catalan and European Portuguese, but also in languages such as English, where there is a strong co-variation between stress and vowel quality.

Therefore, the observations that longer mid-central vowels induce the perception of stress and that duration and intensity cues are sufficient to perceive stress in words made up of open vowels even in languages with a strong co-variation between stress and vowel quality do not support the idea that the degree of co-variation between vowel reduction and stress in a language determines the specific cues that speakers of that language use to perceive stress. Rather, they indicate that the phonetic realization of stress is made up of a cluster of cues with trading relationships.

Results from Catalan illustrate the trading relationships between duration cues and vowel quality, overall intensity, and spectral tilt cues. Duration cues are used in all stimuli, regardless of whether or not they contain vowel quality, overall intensity, or spectral tilt cues in a competing or in an enhancing relationship with duration. In contrast with duration, overall intensity and spectral tilt cues are most apparent in stimuli in which other cues to the stress contrast are ambiguous. These results for Catalan support those for English (Fry, 1955, 1958; Turk and Sawusch, 1996) and Spanish (Ortega-Llebaria *et al.*, 2008) in that duration and intensity cues are in an additive relationship, and, yet, there is an asymmetry in favor of duration. Variations in duration have a stronger effect than variations in intensity in stress perception. However, contrary to Turk and Sawusch's (1996) results, speakers of Catalan were able to perceive stress in stimuli that contained duration cues to stress, but also in those that only had overall intensity or spectral tilt cues to stress, indicating that Catalan speakers used any cue available in the speech signal.

In contrast with overall intensity cues, which were used to perceive stress in both [mama] and [mimi] stimuli, Catalan speakers used spectral tilt cues to perceive stress in [mama] but not in [mimi]. This asymmetry may relate to the spectral tilt manipulations used in this perception experiment, which boosted $F1$, $F2$, and $F3$ amplitudes in [a] but only $F2$ and $F3$ amplitudes in [i]. However, in Spanish (Ortega-Llebaria *et al.*, 2008), where vowels maintain the same quality across stress contexts, speakers did not use either spectral tilt cues to perceive stress in [mama] in spite of performing spectral tilt manipulations identical to those performed in this experiment for Catalan [mama]. Moreover, Catalan speakers in our pilot production experiment produced significant spectral tilt differences in relation to stress only when spectral tilt was measured in stressed [a] and unstressed [ə] without correcting for formant frequency differences. After correcting for formant frequency differences using Fulop *et al.*'s (1998) corrections, there were no clear differences in spectral tilt between these two vowels. Thus, the different use that Catalan and Spanish speakers make of spectral tilt cues to perceive stress in [mama] indicates that, in these languages, spectral tilt changes as a consequence of the formant frequency differences between corresponding vowels rather than in relation to the stress contrast, because only in Catalan stressed [a] reduces to an unstressed [ə] by lowering $F1$ by 200 Hz.

Spectral tilt results for Catalan and Spanish contrast with those obtained for Dutch (Sluijter and van Heuven, 1996; Sluijter *et al.*, 1997), where spectral tilt, together with duration, was a consistent cue to the stress contrast. However, Sluijter *et al.* (1997) examined the perception of spectral tilt in relation to stress only in the vowel [a], which in Dutch reduces to [•] by lowering $F1$ and $F2$ and raising $F3$ (Sluijter and van Heuven, 1996, p. 2481). Consequently, they could not disentangle the effects of vowel reduction in spectral tilt from those of stress.

There is no general agreement cross-linguistically that spectral tilt is a consistent cue to word stress. Although Sluijter (1995) showed that English behaves like Dutch with respect spectral tilt, Campbell and Beckman (1997) could not replicate Sluijter's (1995) results, and Kochanski *et al.* (2005) showed that variations in loudness and duration, but not in spectral tilt or $F0$, were reliable cues to word stress. Other perception studies could not show that spectral tilt is a cue to word stress in Spanish (Ortega-Llebaria *et al.*, 2008)

or to accent in Swedish (Heldner, 2003). In order to disentangle whether spectral tilt is a cue to stress rather than a consequence of the formant frequency differences related to vowel reduction, future research needs to examine vocal fold movement in a set of stressed and unstressed vowels that vary in their qualities. Thus, in Cutler's (2005) (p. 270) words "on this issue, the last word may not yet been spoken."

V. CONCLUSION

Results from Catalan agree with research in other languages that shows that listeners perceive word stress in the absence of pitch accents by relying on a cluster of cues in which duration plays an important role (see, for instance, Fry, 1955, 1958, 1964; Turk and Sawusch, 1996; and Kochanski *et al.*, 2005 for English; Sluijter *et al.*, 1997 for Dutch; and Ortega-Llebaria *et al.*, 2008 for Spanish). They also agree with Lieberman's (1960) finding for English that cues from this cluster trade off, so that in the absence of one cue, speakers rely on others. For example, our results show that in the absence of pitch accents, duration, and vowel reduction cues, listeners perceive stress by relying on overall intensity. Or in the absence of pitch accents, overall intensity, and vowel reduction, speakers rely on duration cues.

Trading relations between duration and vowel reduction were also examined in this experiment. In Catalan, as in English and other languages, vowel reduction implies duration and formant frequency differences between corresponding vowels so that vowels with open qualities are longer than their mid-central counterparts (Lindblom, 1963; Moon and Lindblom, 1994). In spite of this consistent co-variation between duration and vowel quality, Catalan speakers, such as Fry's (1964) English speakers, perceived longer mid-central vowels as more stressed than shorter open vowels, proving that duration differences in relation to stress are not perceived only through vowel quality differences linked to vowel reduction. Thus, cross-linguistic research provides cumulative evidence that word stress is perceived independently of pitch accents by relying on a set of cues with trading effects, so that no single cue, including vowel reduction, is absolutely necessary for the perception of stress.

ACKNOWLEDGMENTS

We thank Mary Beckman and Scott Myers for their insightful comments on earlier versions of this paper. Preliminary results were presented in PaPI 2007. We thank the audience for their comments. Heartfelt thanks to Jane Opie and Michael Mahometa for their valuable help with English and statistics. This research was funded by grants from the Ministerio de Educación y Ciencia-FEDER HUM2006-01758/PHYLLO and by a URAP grant from the College of Liberal Arts, UT-Austin.

Astruch, L. (**2005**). "The intonation of extra-sentential elements in Catalan and English," Ph.D. thesis, University of Cambridge, Cambridge.
Beckman, M. E. (**1986**). *Stress and Non-Stress Accent* (Foris, Dordrecht).
Beckman, M. E., and Edwards, J. (**1994**). "Articulatory evidence for differentiating stress categories," in *Phonological Structure and Phonetic Form. Papers in Laboratory Phonology III*, edited by P. A. Keating (Cambridge University Press, Cambridge), pp. 7–33.
Boersma, P., and Weenink, D. (**2005**). PRAAT: doing phonetics by computer (Version 4.3.01).
Bolinger, D. L. (**1958**). "A theory of pitch accent in English," Word **14**, 109–149.
Campbell, N., and Beckman, M. E. (**1997**). "Stress, prominence and spectral tilt," *Intonation: Theory, Models and Applications*, *Proceedings of an ESCA Workshop*, Athens, Greece, 18–20 September 1997, edited by A. Botinis, G. Kouroupetroglou, and G. Carayiannis (Department of Informatics, University of Athens, Greece), pp. 67–70.
Chen, T., and Massaro, D. (**2004**). "Mandarin speech perception by ear and eye follows a universal principle," Percept. Psychophys. **66**, 820–836.
Cutler, A. (**2005**). "Lexical stress," in *The Handbook of Speech Perception*, edited by D. B. Pisoni and R. E. Remez (Blackwell, Oxford), pp. 264–289.
de Jong, K. (**1991**). "The oral articulation of English stress accent," Ph.D. thesis, The Ohio State University, Columbus, OH.
Delgado Martins, M. R. (**1975**). "Perception of Portuguese stress," Proceedings of the VIII International Congress of Phonetic Sciences, Leeds, United Kingdom.
Fry, D. B. (**1955**). "Duration and intensity as physical correlates of linguistic stress," J. Acoust. Soc. Am. **27**, 765–768.
Fry, D. B. (**1958**). "Experiments in the perception of stress," Lang Speech **1**, 126–152.
Fry, D. B. (**1964**). "The dependence of stress judgments on vowel formant structure," Proceedings of the 5th International Congress in Phonetic Sciences, Münster, Germany, edited by S. Karger (Basel, New York), pp. 306–311.
Fulop, S. A., Kari, E., and Ladefoged, P. (**1998**). "An acoustic study of the tongue root contrast in Degema vowels," Phonetica **55**, 80–98.
Heldner, M. (**2003**). "On the reliability of overall intensity and spectral emphasis as acoustic correlates of focal accents in Swedish," J. Phonetics **31**, 39–62.
Herrick, D. (**2003**). "An acoustic analysis of phonological vowel reduction in six varieties of Catalan," Ph.D. thesis, University of California, Santa Cruz, CA.
Huss, V. (**1978**). "English word stress in the post-nuclear position," Phonetica **35**, 86–105.
Kochanski, G., Grabe, E., Coleman, J., and Rosner, B. (**2005**). "Loudness predicts prominence: Fundamental frequency lends little," J. Acoust. Soc. Am. **118**, 1038–1054.
Lieberman, P. (**1960**). "Some acoustic correlates of word stress in American English," J. Acoust. Soc. Am. **32**, 451–454.
Lindblom, B. (**1963**). "Spectrographic study of vowel reduction," J. Acoust. Soc. Am. **35**, 1773–1781.
Massaro, D., Cohen, M., Gesi, A., Heredia, M., and Tsuzaki, M. (**1993**). "Bimodal speech perception: An examination across languages," J. Phonetics **21**, 445–478.
Moon, S., and Lindblom, B. (**1994**). "Interaction between duration, context, and speaking style in English stressed vowels," J. Acoust. Soc. Am. **96**, 40–55.
Ortega-Llebaria, M., Prieto, P., and Vanrell, M. M. (**2008**). "Perceptual evidence for direct acoustic correlates of stress in Spanish," Proceedings of the 16th International Congress of Phonetic Sciences, Saarbrücken, Germany, 6–10 August 2007, edited by J. Trouvain and W. J. Barry, pp. 1121–1124.
Pierrehumbert, J. B. (**1980**). "The phonetics and phonology of English intonation," Ph.D. thesis, Massachusetts Institute of Technology, Cambridge, MA.
Recasens, D. (**1986**). *Estudis de Fonètica Experimental del Català Oriental Central (Experimental Phonetic Studies on East-Central Catalan)* (Publicacions de l'Abadia de Montserrat, Barcelona, Spain).
Recasens, D., and Espinosa, A. (**2006**). "Dispersion and variability of Catalan vowels," Speech Commun. **48**, 645–666.
Sluijter, A. M. C. (**1995**). *Phonetic Correlates of Stress and Accent* (Holland Academic Graphics, The Hague).
Sluijter, A. M. C., and van Heuven, V. (**1996**). "Spectral balance as an acoustic correlate of linguistic stress," J. Acoust. Soc. Am. **100**, 2471–2485.
Sluijter, A. M. C., van Heuven, V., and Pacilly, J. A. (**1997**). "Spectral balance as a cue in the perception of linguistic stress," J. Acoust. Soc. Am. **101**, 503–513.
Turk, A., and Sawusch, J. (**1996**). "The processing of duration and intensity cues to prominence," J. Acoust. Soc. Am. **99**, 3782–3790.
Wheeler, M. W. (**2005**). *The Phonology of Catalan* (Oxford University Press, Oxford, UK).