

Weakening of Intervocalic /s/ in the Nijmegen Corpus of Casual Spanish

Francisco Torreira^a Mirjam Ernestus^{a,b}

^aMax Planck Institute for Psycholinguistics and ^bCLS, Radboud Universiteit, Nijmegen, The Netherlands

Abstract

This study describes the weakening of intervocalic /s/ in the Nijmegen Corpus of Casual Spanish and investigates the role of several potential conditioning factors, including morphological, lexical and probabilistic variables. Three acoustic parameters were examined: voicing, the difference in high-band (4–8 kHz) intensity between /s/ and the following vowel, and the duration of the dip in low-band (0–1.5 kHz) intensity in /s/. Over a third of the /s/ consonants exhibited uninterrupted voicing. All three parameters were highly sensitive to speech rate and, less consistently, to prosodic factors such as word position and stress. We also found that /s/ suffixes in redundant morphosyntactic contexts were more prone to weakening than other word-final /s/ segments. Several high-frequency words were particularly prone to weakening, but no general probabilistic effects were observed for factors claimed to favor reduction such as word frequency, contextual predictability or grammatical class. These findings demonstrate the complex variability characteristic of reduction phenomena in spontaneous speech.

Copyright © 2012 S. Karger AG, Basel

1 Introduction

1.1 Speech Reduction and Goal of the Present Study

The reduction of consonants and vowels is pervasive in spontaneous conversational speech, as attested in a rising number of corpus studies [Kohler, 2000; Ernestus, 2000; Johnson, 2004; Adda-Decker et al., 2005; Schuppler et al., 2011; Adda-Decker and Snoeren, 2011; Bürki et al., 2011]. In spite of the increasing interest in speech reduction phenomena, much remains to be explored about their phonetic characteristics and conditioning factors. For instance, many corpus studies have focused on the effects of probabilistic, morphological and lexical factors on word and segmental durations, but studies examining the effects of such factors on other phonetic parameters and on specific reduction phenomena are scarce.

The present work describes the weakening of intervocalic /s/ in Spanish, an almost undocumented phenomenon despite its high frequency in conversational speech, and

investigates its conditioning factors. During the initial inspections of the Nijmegen Corpus of Casual Spanish (NCCSp from now on) [Torreira and Ernestus, 2010a], a recently created corpus of Madrid Spanish, we observed that intervocalic /s/ consonants, although always realized as sibilants, often exhibit short durations, uninterrupted voicing and decreased frication noise, resulting in a percept of consonantal weakening, i.e. such /s/ consonants seem less acoustically salient between vowels than longer, voiceless /s/ consonants exhibiting stronger frication (see subsection 1.2 below). In contrast to the aspiration and deletion of syllable-final /s/ ([s] > [h] > Ø) in coastal Latin American and southern Spanish dialects, a reduction process which has been the object of numerous studies (see extensive bibliography in Hernández-Campoy and Trudgill [2002]), the sort of /s/ reduction that we investigate here has only been described impressionistically in the philological literature for a few rural dialects of Castilian Spanish [Torreblanca, 1986]. This sort of /s/ reduction does not involve a complete loss of the oral constriction as in /s/ aspiration, and is referred to as /s/ weakening from now on for the sake of simplicity. In the present article, we describe it as it occurs in the NCCSp (see subsection 1.6 below for a description of this corpus), using quantitative acoustic phonetic methods.

In study 1, we provide an initial overview of the phenomenon by estimating the frequency with which it occurs, and by examining the extent to which it is influenced by factors known to condition coarticulatory and reduction phenomena such as speech rate, stress, position in the word, number of syllables in the word and the segmental context. In studies 2 and 3, we explore the role of morphosyntactic, probabilistic and lexical factors after controlling for the relevant conditioning factors identified in study 1. Since /s/ is the most frequent consonant in Spanish [Moreno Sandoval et al., 2008], occurs in a wide variety of function and content words, and often carries morphosyntactic information as a plural marker or verbal suffix (e.g. *las casas* ‘the houses’; *vas* ‘you go’), /s/ weakening provides an ideal case for investigating the roles of these variables in segmental reduction.

So far, most studies investigating potential conditioning factors of reduction phenomena (e.g. content vs. function word distinction, lexical frequency and the probability of words given surrounding words) have used global measures of reduction such as word, morpheme or segment durations, on the basis that they should capture most kinds of reduction-related phonetic and phonological processes in a data set (e.g. deletions, vowel reduction). In the present study we adopt a different strategy and investigate whether the effects of such variables can also be observed on more fine-grained acoustic measures in a specific reduction phenomenon such as /s/ weakening. In all three studies, we examine three acoustic parameters: the duration of the dip in low-band (0–1.5 kHz) intensity corresponding to /s/, the occurrence of uninterrupted voicing, and the difference in high-band (4–8 kHz) intensity between /s/ and the following vowel.

1.2 Realizations of Intervocalic /s/ in the NCCSp

This subsection illustrates the variability in the realization of intervocalic /s/ present in the NCCSp with examples extracted from this corpus. Figure 1¹ shows canonical and slightly weakened realizations of intervocalic /s/. The example at the

¹ Sound files corresponding to figures 1–4 can be accessed online (for all online supplementary material, see www.karger.com/doi/10.1159/000343635).

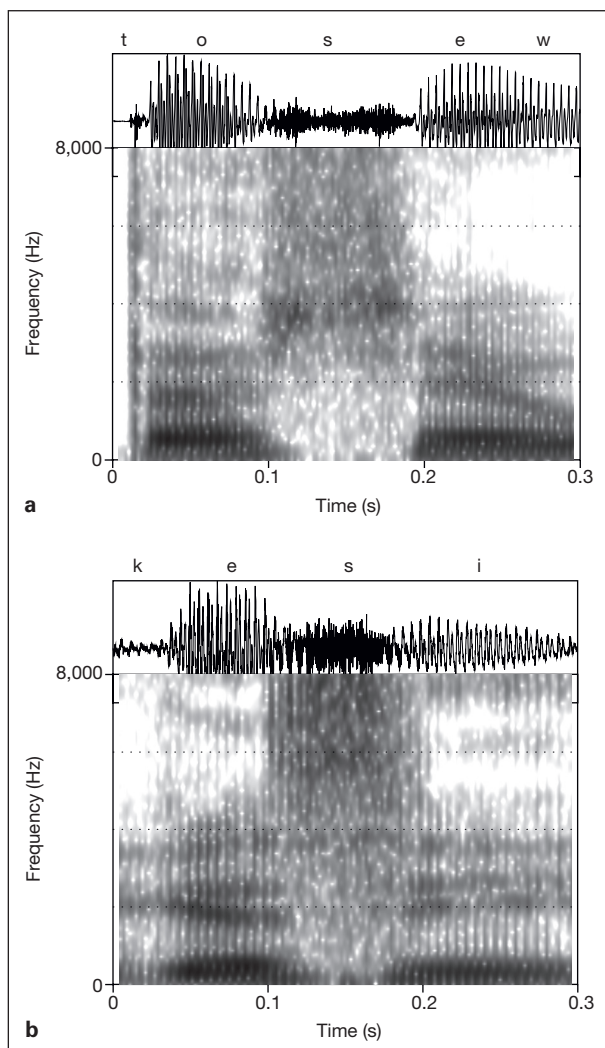


Fig. 1. Waveforms and spectrograms of utterances containing canonical (top) and slightly weakened (bottom) realizations of intervocalic /s/. **a** The example has been extracted from the utterance *ochocientos euros* ‘eight hundred euros’. **b** The second example corresponds to the utterance *que sí* ‘definitely yes’.

top illustrates a canonical pronunciation containing the acoustic features expected for an intervocalic /s/. Voicing is interrupted for around 80 ms, as shown by the lack of a voicing bar and periodic energy in the time region of the spectrogram corresponding to the fricative consonant. In this time region, energy drops in the lower part of the spectrum, reflecting the formation of a front oral constriction, and sharply increases above 3.5 kHz, reflecting the creation of turbulence at the constriction. It is this prominent concentration of aperiodic energy in mid-high regions of the spectrum that gives [s] its characteristic sibilant quality [Shadle and Scully, 1995; Stevens, 1998]. The bottom example in figure 1 displays a similar pattern in the mid-high frequencies. However, in this example voicing is present throughout much of the realization of the fricative and gradually disappears towards its mid-point, before

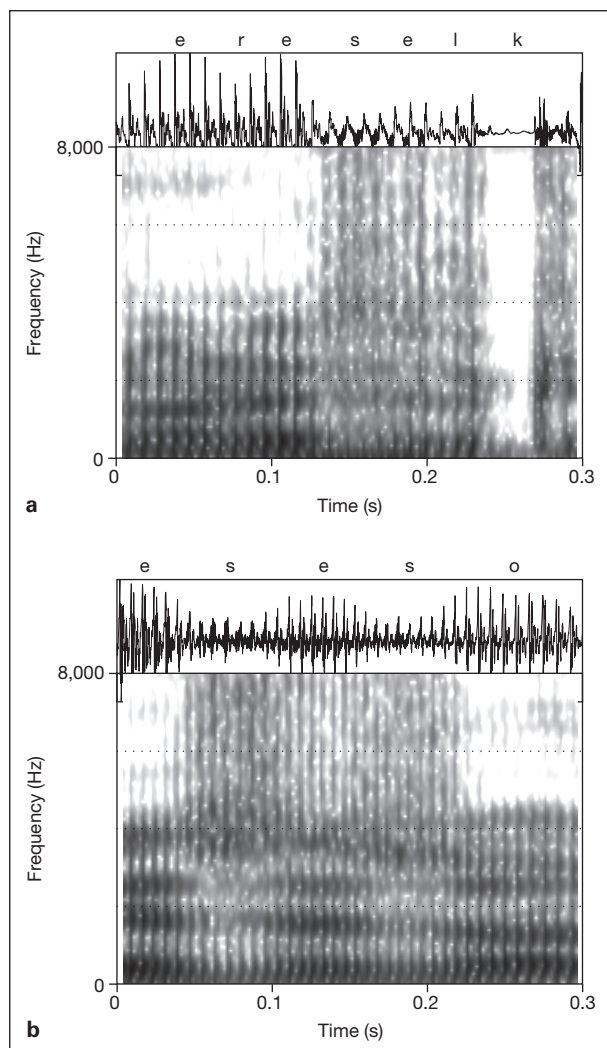


Fig. 2. Waveforms and spectrograms of utterances containing weakened intervocalic /s/. **a** The example has been extracted from the utterance *tú no eres el que* [. . .] ‘you are not the one who [. . .]’. **b** The second example has been extracted from the utterance *es eso* ‘that’s it’.

picking up as the transition into the upcoming vowel approaches. Moreover, one can also see a less sharp decrease in low-frequency energy than in the top example, resulting perhaps from a less narrow oral constriction and the presence of voicing.

The examples in figure 2 show further weakened realizations of intervocalic /s/. Frication noise in the mid-high frequencies is less prominent than in the examples in figure 1, and in both cases voicing continues uninterrupted throughout the [s] consonants with no visible signs of decay. Moreover, the duration of the consonants, as estimated informally from the amplitude envelope in the waveforms and the occurrence of aperiodic energy in the spectrograms, appears to be significantly shorter than that of the examples in figure 1. The examples in figure 3 illustrate cases of extreme weakening. Frication noise in these examples is considerably weaker than in

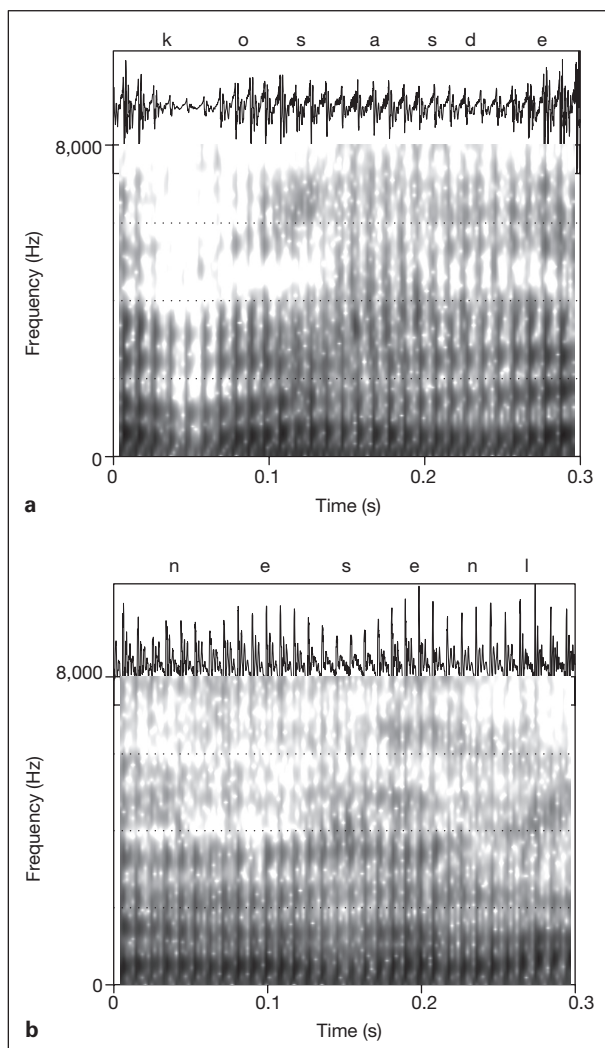


Fig. 3. Waveforms and spectrograms of utterances containing extremely weakened realizations of intervocalic /s/. **a** The example at the top has been extracted from the utterance [. . .] *mazo cosas de* [. . .] ‘[. . .] a lot of things about [. . .]’. **b** The second example has been extracted from the utterance *corrupción es en la* [. . .] ‘[. . .] corruption is in the [. . .]’.

the previous ones, to the point that it is hardly visible. Voicing, as in the examples in figure 2, continues uninterrupted throughout the vowel-/s/-vowel sequences.

These examples make it apparent that the familiar correlates of a voiceless fricative (aperiodic energy in the high frequencies and lack of periodicity throughout the consonant) are not always present in our data, and that, if present, they may be very weak. For this reason, segmental durations and related measures, such as the proportion of the fricative containing voicing, cannot be made in a consistent way across the tokens in our data. The study of /s/ weakening in spontaneous Spanish must, therefore, be based on other acoustic measures. Section 2.1.1 presents the measures that we employed, namely the occurrence of uninterrupted voicing, the difference in high-band (4–8 kHz) intensity between [s] and the following

vowel, and the duration of the dip in low-band (0–1.5 kHz) intensity corresponding to [s].

1.3 Possible Articulatory Mechanisms behind /s/ Weakening

The voicing and weak frication in Spanish intervocalic /s/ illustrated in the previous subsection may result from different mechanisms. Given that the segments surrounding /s/ require the production of voicing and that the intervening /s/ requires interrupting it, it is plausible that voicing in intervocalic /s/ is a coarticulatory phenomenon. Generally speaking, voicing in intervocalic /s/ is likely to arise from the conflicting demands of adjacent segments on glottal articulation when the glottis must adopt opposite configurations [Hoole, 1999].

The high-frequency frication characteristic of [s] has its source in a jet of air passing through a narrow coronal constriction. This jet of air hits the incisors, creating an additional turbulent noise source, and enters a small front cavity between the constriction and the lips with a high-frequency resonance [Shadle and Scully, 1995; Stevens, 1998]. The weakened frication observed in many of our /s/ tokens (fig. 2, 3) can be attributed to at least two mechanisms. First, the supraglottal constriction may not be narrow enough for the creation of turbulence. This could be caused by the contextual influence of adjacent vowel sounds with constriction targets specifying an open vocal tract. Second, air volume velocity at the supraglottal constriction may not be sufficient for the creation of strong turbulent noise. This could be the consequence of the glottal coarticulation discussed in the previous paragraph. During the production of intervocalic [s], the glottis may not always reach a fully abducted position under the coarticulatory influence of adjacent vowels, impeding airflow and hindering the creation of turbulence at the downstream supraglottal constriction [Ohala and Solé, 2010].

In the present work we do not intend to investigate the exact articulatory mechanisms leading to the frequent weakening of intervocalic /s/ in Spanish. To do this we would need articulatory and aerodynamic data collected under highly controlled conditions. In study 1, where we present an exploratory overview of the phenomenon, we will just examine the extent to which /s/ weakening is sensitive to factors known to condition coarticulation, and more generally, speech reduction. We expect /s/ weakening to be favored in temporally reduced realizations, including those produced at higher speech rates and in prosodically weaker positions (i.e. unstressed position, non-word-initial position, and longer words).

1.4 Morphological Effects on Speech Reduction

Recent studies indicate that reduction does not affect all words and morphemes to the same degree. For instance, it has been claimed that segments that carry semantic information by themselves, such as 1-segment suffixes, tend to be less reduced than segments that are part of a stem. For example, several studies have reported higher rates of deletion for English word-final /t/ and /d/ consonants when they are part of a stem (as in *mist*) than when they are a verbal suffix (as in *missed* [Guy, 1991; Bybee, 2002; Labov, 2004]). Such differences could arise either from morphological effects on phonological encoding percolating down to articulation, or, indirectly, via mechanisms such as those proposed by van Son and Pols [2003], Pluymaekers et al. [2005a, b], and Aylett and Turk [2004], in which the speaker modulates articulatory effort according to the informativeness of the speech signal.

Regarding /s/ weakening, it should be noted that word-final /s/ in Spanish can occur as a plural marker in determiners, nouns and adjectives (e.g. *las niñas bonitas* ‘the pretty girls’), as a verbal suffix (e.g. *come* ‘she eats’ versus *comes* ‘you eat’), and as part of a larger morpheme, either a root or a suffix (e.g. *miércoles* ‘Wednesday’). In line with the morphological effects on final /t/ and /d/ deletion in English mentioned above, it could be hypothesized that, in Spanish, /s/ suffixes (e.g. *comes* ‘you eat’) are less prone to reduction than /s/ segments that are part of a larger morpheme (e.g. *miércoles* ‘Wednesday’). However, this hypothesis has not been confirmed by several variationist studies on the realization of /s/ in several Latin American Spanish dialects [Poplack, 1980; Hundley, 1987; Lafford, 1989]. Instead, several of these studies have found the opposite: word-final /s/ is aspirated and deleted more often when it is a suffix than when it is part of a stem. A possible explanation for this is that /s/ suffixes in Spanish are often redundant based on the morphological, semantic or syntactic information in the utterance (e.g. *cuatro cosas* ‘four things’). Poplack [1980], for Puerto Rican Spanish, and Hundley [1987], for Peruvian Spanish, found support for this explanation in that /s/ suffixes in these dialects are more likely to be aspirated or deleted when they are morphosyntactically redundant given the context. Similarly, Erker [2010] found that the duration of sibilant realizations of /s/ in Dominican Spanish tends to decrease when the target word is preceded by other words with -s plural suffixes in the same noun phrase.

In study 2, we test if the weakening of word-final /s/ in spontaneous Madrid Spanish is affected by morphological factors such as the ones discussed above. In particular, we will examine if /s/ weakening is affected by morphosyntactic contextual redundancy or if it is sensitive to morphological structure regardless of contextual redundancy.

1.5 Probabilistic and Lexical Effects on Speech Reduction

A series of studies on English and Dutch have shown that there is a tendency towards more reduction in highly frequent and predictable words, in function words compared to content words, and in suffixes highly predictable from their stems [Bell et al., 2003, 2009; Aylett and Turk, 2004, 2006; Pluymaekers et al., 2005a, b; Ernestus et al., 2006; Baker et al., 2011]. These studies have examined a range of dependent variables, such as segmental transcriptions of full versus reduced forms [Bell et al., 2003], word duration [Bell et al., 2009; Baker et al., 2011], syllable duration [Aylett and Turk, 2004], vowel duration and spectral characteristics [Aylett and Turk, 2006], suffix duration and number of realized segments as opposed to underlying segments [Pluymaekers et al., 2005a, b], and segmental duration and voicing in consonant clusters [Ernestus et al., 2006]. Regarding Spanish, word frequency has been claimed to favor the aspiration and deletion of /s/ in several Latin American dialects [Brown and Torres Cacoullós, 2003; Minnick Fox, 2006; File-Muriel, 2007; Brown, 2009]. These probabilistic and lexical effects on articulation have been interpreted as the consequence of different speech production mechanisms. These include the repetition of frequent articulatory routines [Bybee, 2001], a modulation of articulatory effort sensitive to informativeness [Aylett and Turk, 2004; Pluymaekers et al., 2005a, b], and lexical access, which is claimed to operate differently for function words and content words [Bell et al., 2009].

On the other hand, several studies suggest that probabilistic and lexical effects on speech reduction may be more complex than is often claimed in the literature. Recent work on the realization of /s/ in Caleño Spanish [File-Muriel and Brown, 2011] report longer [s] durations the higher the lexical frequency, and the higher

spectral center of gravity values and percentages of voicelessness, the higher the word bigram frequency of the target word and its following word (see regression coefficients in their tables 1 and 2). Similarly, Kuperman et al. [2007] found that Dutch -s- interfixes were longer the more probable they were given the compound in which they occurred and its constituents. Finally, Torreira and Ernestus [2010b] investigated phrase-medial vowel devoicing in spontaneous French, and found that devoicing was as common in content words as in high-frequency function words. Since Spanish /s/ occurs in function as well as in content words representing a wide range of lexical frequencies, it offers an ideal case for further testing the role of these factors on speech reduction. Study 3 examines the influence of these factors on /s/ weakening.

1.6 *The Nijmegen Corpus of Casual Spanish*

All of the speech data used in the studies presented in this paper come from the NCCSp. This recently created corpus consists of recordings of 52 Spanish speakers (27 female and 25 male) engaged in casual conversations with friends. The corpus, which features around 30 h of high-quality audio and video recordings, has been orthographically annotated by professional transcribers at Verbio Speech Technologies S.L. The transcriptions of the corpus contain around 340,000 word tokens and 16,500 word types (distinct lexical items).

All speakers were university students in their late teens or twenties and had been raised in the Madrid region. Although they were of similar age and educational background, it is apparent from the contents of the conversations that they varied considerably in other social dimensions (e.g. socioeconomic background). Moreover, inspection of questionnaires filled in by the speakers after each recording session revealed that most of them had at least one parent born outside of the region. Our speakers therefore reflect to a certain extent the social diversity present in an urban center like Madrid.

The NCCSp was collected in a sound-attenuated room at the Universidad Politécnica in Madrid during March 2009. In each recording session, which lasted around 90 min, we recorded a group of 3 friends of the same sex: 2 naive speakers and 1 confederate. The role of the confederate was to recruit the other two participants (the naive speakers), and to make the conversation livelier whenever it approached a dead end. Eight naive speakers took part in subsequent recordings as confederates. All participants wore a Samson QV head-mounted unidirectional microphone. The microphones were placed at an average distance of 5 cm from the left corner of the speakers' lips. The naive speakers were captured in separate channels of a stereo signal with an Edirol R-09 solid-state recorder. The confederate was directly recorded on a computer via a dedicated sound card. The sampling rate used was 44.1 kHz, while quantization was set to 16 bits.

The recording procedure used for the NCCSp was similar to that previously used for the Nijmegen Corpus of Casual French [Torreira et al., 2010]. Each recording session consisted of three parts. In part 1, we told the participants that the confederate's microphone did not work properly and asked the confederate to leave the room. While waiting, the naive speakers spontaneously engaged in free conversation, which was recorded unbeknownst to them. This part lasted around 20 min on average. Part 2 consisted of free conversation between the confederate and their friends, which lasted around 35 min. Part 3 required participants to choose three questions from a list of general interest questions, and to negotiate and provide group answers to these questions

in around 35 min. At the end of the recording, we disclosed our procedures to the participants and paid EUR 30 to each of the speakers and EUR 45 to the confederate as a compensation for their time. All of the participants signed a consent form agreeing to the use of the recordings for academic and scientific purposes. More detailed information about the corpus can be found online at <http://www.mirjamernestus.nl/Ernestus/NCCSp>, and in Torreira and Ernestus [2010a].

1.7 Outline of this Article

Sections 2–4 present three corpus studies carried out to investigate the occurrence of /s/ weakening in spontaneous Spanish. Study 1 provides an exploratory overview of this reduction phenomenon. We estimate its frequency and to what extent it is sensitive to several factors known to condition coarticulation and speech reduction in general. In study 2, we examine the potential role of morphosyntactic factors on /s/ weakening. We investigate if word-final /s/ is more weakened when it is part of a stem than when it is a suffix, and, within the group of suffixes, if grammatically redundant suffixes show a bias towards more weakening. Finally, study 3 investigates the potential role of lexical and probabilistic variables on /s/ weakening, including grammatical class (content vs. function words), lexical frequency and contextual predictability, and word-specific effects.

2 Study 1: Exploratory Overview of /s/ Weakening

2.1 Method

2.1.1 Materials

As a first step, we randomly extracted from the NCCSp 1,100 intervocalic [s] consonants not disrupted by laughter, disfluencies or intrusive overlapping speech. Then we sampled additional tokens meeting the same conditions until every speaker in the corpus was represented by at least 20 tokens. This data set contained 1,257 tokens and 593 word types. The average number of word tokens by word type was 3.3, with a standard deviation of 9.9 and a range extending from 1 to 93.

In order to roughly locate vowel-/s/-vowel sequences in the corpus, a forced segmental alignment of the NCCSp was performed with HTK (Hidden Markov Model Toolkit) automatic speech recognition software [Young et al., 2002]. Using the orthographic transcription of the corpus and a dictionary of canonical pronunciations, a monophone system consisting of 25 phone categories corresponding to Spanish vowels and consonants and a silence phone were trained on the NCCSp. Each monophone was represented by a 3-state left-to-right hidden Markov model. The frame length of each state was 10 ms. Therefore, phones in the alignment were always assigned a duration of minimally 30 ms. This alignment was also used to estimate speech rate, an important variable in our analyses (see section 2.1.4).

As explained in section 1.2, the usual correlates of a voiceless fricative (aperiodic energy in the high frequencies and lack of periodicity) were present in such varying degrees across tokens that they could not be used in a consistent way to determine the start and end of [s] consonants. For this reason, we did not perform temporal measurements using traditional segmentation procedures. Instead, we based our measurements on the energy profile of the vowel-[s]-vowel sequence. The following measurements were made.

2.1.1.1 Low-Band (0–1.5 kHz) Intensity Dip Duration. During the initial inspection of the data, we noticed that all tokens exhibited a dip in energy in the lower frequencies in the part of the signal corresponding to the articulation of /s/. Since this dip can be said to result from the formation of an oral constriction, its duration may prove a useful measure with which we can estimate the temporal reduction of /s/ consonants. This duration measure cannot be considered at the same level as a *segmental* duration, and, accordingly, it cannot be used to compute related measures, such as the amount of voicing within the [s] segment. This is not a technical limitation of our work, but a direct consequence of the fact that segments in spontaneous speech often do not present clear boundaries and exhibit more variable realizations than in controlled speech.

In order to measure the duration of the low-band intensity dip in the vowel-[s]-vowel sequence, we used the following procedure illustrated by the solid curves and dotted vertical lines in the intensity panels of figure 4. First, we low-pass filtered the signals below 1.5 kHz. Second, we placed marks at the midpoints of the automatically segmented vowels, and then checked if the interval between them included the intensity velocity minima and maxima in the vowel-to-[s] and [s]-to-vowel transitions. In those cases in which this was not the case, we manually corrected the placement of the marks (but not the automatic segmentation). Third, we calculated the intensity contours of the low-passed signals using a 30-ms window and a time step of 8 ms, and differentiated these contours by first smoothing them via cubic interpolation and then calculating local differences in intensity at intervals of 1 ms. Finally, we automatically identified the locations of the intensity velocity minimum and maximum in the vowel-to-[s] and [s]-to-vowel transitions. These points, located where intensity below 1.5 kHz was decreasing and increasing most rapidly, were considered as the start and the end of the dip. The low-band intensity dip duration was then derived from these time points².

2.1.1.2 Voicing. The occurrence of voicing in the low-band intensity dip was estimated with the pitch detection algorithm available in Praat set to default parameters. We observed that most tokens showed either uninterrupted voicing, or a stretch of voicelessness encompassing the intensity dip. For this reason, the amount of voicing within the dip cannot be used to study gradient changes in voicing duration within [s]. We therefore created a categorical variable encoding the presence of uninterrupted voicing: each /s/ token was labeled as *voiced* when no interruption of pitch periods was observed throughout the dip in low-frequency intensity, and as *voiceless* otherwise. Unvoiced /s/ tokens can therefore present variable degrees of voicing throughout their realization (in parts of the signal other than the low-frequency intensity dip), resulting from variable degrees of weakening at the glottal level. Voiced tokens, on the other hand, can all be regarded as instances of substantial weakening.

2.1.1.3 High-Band (4–8 kHz) Intensity Difference. Clearly articulated /s/ consonants typically exhibit more acoustic energy than vowels in the high frequencies,

² An anonymous reviewer notes that the variable presence of intrusive voicing throughout the [s] consonant may have had undesired effects on our duration measure, for instance by producing longer dip durations in voiced [s] due to a shallower dip. This did not occur in our data (see section 2.2.1. below). Moreover, using a 0.5- to 1.5-kHz band aimed at counteracting the effects of voicing excluded the energy of the first vowel formant in numerous vowel-[s] transitions, leading to an inconsistent placement of low-band intensity velocity minima and maxima. For these reasons, we decided to use a low-frequency band including both voicing and F₁ in all cases.

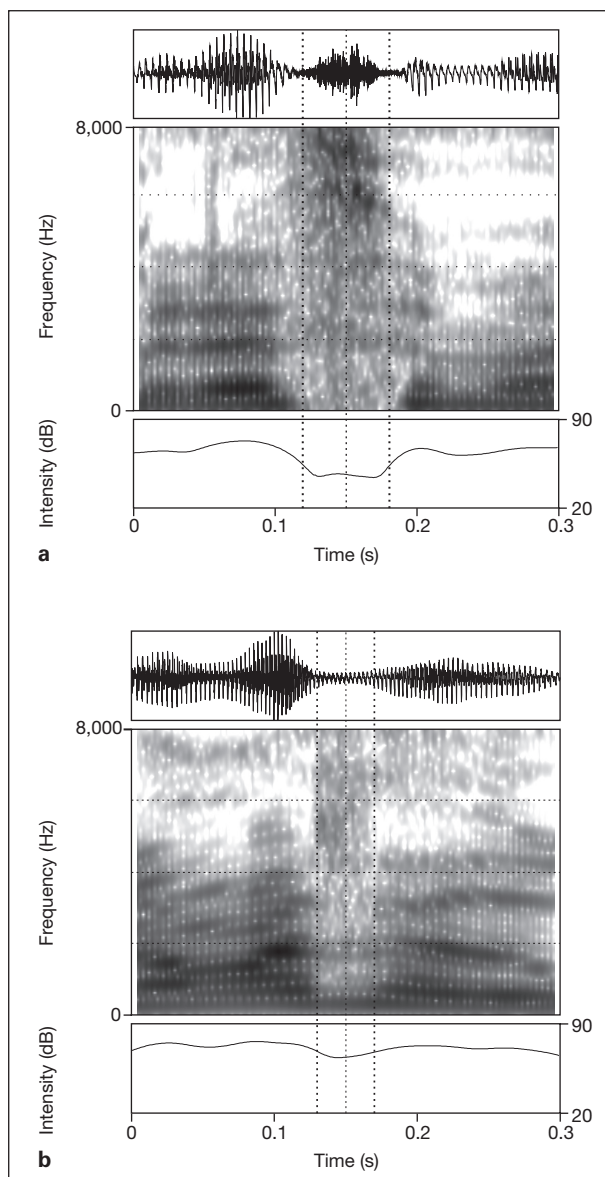


Fig. 4. Waveforms, spectrograms, and intensity contours of examples of clearly articulated (**a**) and weakened (**b**) /s/ consonants. The solid curves in the intensity panels correspond to the intensity contours of signals low-passed filtered at 1.5 kHz. The onsets, offsets and midpoints of the low-band (0–1.5 kHz) intensity dip are represented by vertical dotted lines (see text for more details on how these time points were determined).

resulting mainly from the presence of aperiodic energy. In order to estimate the strength of the frication noise characteristic of each /s/ token, we defined two 30-ms windows, one centered at the midpoint of the low-band intensity dip, and the other one left-aligned immediately after it (corresponding roughly to the initial part of the following vowel). Then we calculated the difference in intensity in a 4 to 8-kHz band between the intensity peak in the low-band intensity dip window and the lowest intensity value

in the window at the start of the following vowel³. During the initial inspection of the data, this difference appeared to be a better correlate of the intensity of high-frequency frication than more widely used measures such as the spectral center of gravity, which we decided not to use in our analyses. It is referred to as high-band intensity difference from now on. Realizations of /s/ with strong frication noise should exhibit high values of this measure, as opposed to consonants with weak frication.

2.1.2 Predictors

For each intervocalic /s/ token, we coded a number of variables as potential conditioning factors of the measurements of /s/ weakening that we performed (low-band intensity dip duration, voicing and high-band intensity difference). These variables were chosen on the basis of common knowledge about the conditioning factors of coarticulatory and reduction phenomena, and of the ease with which they could be coded automatically. These included speech rate, the immediate segmental context in which /s/ occurred (preceding and following vowel), the position of /s/ in the word (initial vs. medial vs. final), the number of syllables in the word, and whether /s/ was the onset of a stressed syllable (as opposed to that of an unstressed syllable). We also included speaker gender in our predictors, since this variable could potentially affect voicing and spectral measures due to physiological differences between females and males.

In order to estimate speech rate in the utterances in which target /s/ consonants were produced, we used the following procedure. First, we calculated the mean and standard deviation of the log duration of every phone category (consonant or vowel type) in the forced alignment of the NCCSp. Then we computed the log duration of the three phones preceding and following, but not including the target /s/ consonant, and then calculated how many standard deviations each phone log duration was from the mean log duration of its phone category. Our estimate of speech rate was defined as the average value of these deviations. Negative values of this estimate were characteristic of fast speech rates (i.e. as a consequence of phone durations shorter than average), while positive values were characteristic of slow speech rates.

We investigated the role of segmental context by coding the identity of the surrounding vowels. These segments are referred to as S_{i-1} , S_{i+1} , where S_i is the target /s/ consonant. Table 1 shows the number of tokens for each level of these variables.

As for lexical stress, we coded if /s/ was the onset of a stressed syllable or not. In order to determine whether a syllable was stressed, we followed the guidelines provided in Hualde [2009]. In the case of word-final /s/, we assume that resyllabification always takes place (e.g. *los otros* is pronounced as [lo.so.tros] ‘the others’). Word-final /s/ therefore takes the stress value of the first syllable of upcoming words. This was the case for 115 out of the 362 word-final /s/ consonants in the data set. It should be noted that if our resyllabification assumption is not valid, interactions between position in word and stress should be observed in our statistical analyses. Following this procedure, 635 /s/ consonants were coded as stressed, and 622 as unstressed.

³ The lowest intensity value was chosen instead of the highest value because the latter often corresponded to the transition from the [s] consonant and was highly correlated with the intensity level within the consonant. On the other hand, the lowest point tended to correspond to a low-intensity plateau characteristic of the vowels in our data.

Table 1. Number of tokens for each segment type preceding and following the target /s/ (S_{i-1} and S_{i+1})

	/a/	/e/	/i/	/o/	/u/
S_{i-1}	272 (21.6%)	468 (37.2%)	102 (8.1%)	398 (31.6%)	17 (1.3%)
S_{i+1}	334 (26.5%)	403 (32.1%)	247 (19.6%)	198 (15.7%)	75 (5.9%)

2.1.3 Statistical Analysis

In all analyses, we use mixed-effects linear regression with contrast coding as implemented in the *lmer* function in R [Bates and Sarkar, 2006]. For the occurrence of voicing throughout [s], a binary dependent variable (voiceless vs. voiced), we use logistic mixed-effects linear regression. Speaker and word type are modeled as random factors. The other predictors (e.g. speech rate, position in word, number of syllables, preceding and following vowel) are treated as fixed predictors. Only statistically significant predictors are retained in the reported models. Similarly, only statistically significant interactions between predictors are retained. Since determining the number of degrees of freedom in mixed-effects models is a controversial issue, p values are obtained through Monte Carlo Markov Chain simulations [Baayen, 2008].

2.2 Results

2.2.1 Voicing

Over a third (34%, $n = 434$) of the randomly extracted vowel-/s/-vowel sequences contained uninterrupted voicing. Voicing of intervocalic /s/ in casual Spanish therefore appears to be a frequent phenomenon. In order to investigate the different sources of variability behind the voicing of intervocalic /s/, we fitted a mixed-effects regression model considering the variables in section 2.1.2 (i.e. speech rate, preceding and following vowel, position in the word, number of syllables in the word, stress) as fixed predictors, and speaker and word type as random factors. Low-band intensity dip duration, which is the dependent variable in section 2.2.3, was added to the other predictors in the model because we wanted to investigate if voicing co-occurred with the temporal reduction of the consonant.

In order to facilitate the interpretation of regression coefficients, the continuous predictors speech rate and low-band intensity dip duration were standardized into z-scores. Since these continuous predictors were moderately correlated with each other ($r = 0.33$), we orthogonalized these two variables by replacing low-band intensity dip duration with the residuals of a linear model in which low-band intensity dip duration was predicted by speech rate. This variable captures the information in low-band intensity dip duration that cannot be attributed to its linear relationship with speech rate, and allows us to assess the role of low-band intensity dip duration independently from that of speech rate. Since low-band intensity dip duration appeared to be slightly associated with position in the word (word-final /s/ consonants tend to be shorter, see section 2.2.3 below), we decided to run, in addition to our regression models, an ANOVA test in which the role of word position was assessed sequentially after controlling for the effects of duration.

Our mixed-effect model revealed a standard deviation of 0.88 for the random factor speaker, indicating that the average difference in the occurrence of voicing

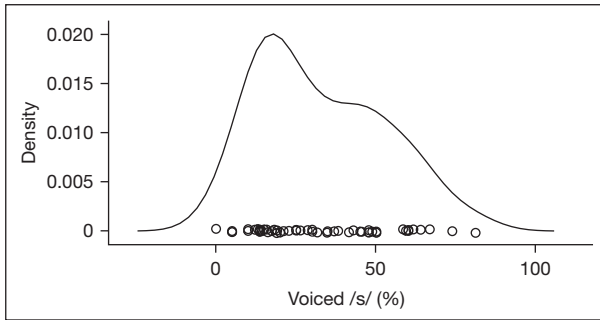


Fig. 5. Kernel density plot of percentages of voicing of intervocalic /s/ per speaker. Each speaker is represented by a circle. Larger areas under the density curve correspond to greater concentrations of circles along the x-axis.

among speakers was approximately 22%. This interspeaker variability is illustrated in the kernel density plot in figure 5, which shows the distribution of percentages of voicing for each speaker. This figure shows that most speakers in our data had a voicing rate below 20%, including a few who rarely voiced their /s/ consonants, and one who never did. High rates of voicing were not uncommon in the data either, with 8 of the 52 speakers exhibiting voicing rates over 50%, and reaching 80% for 2 speakers. The random factor word was not significant.

As for the fixed effects, we found that only speech rate, low-band intensity dip duration and word position were statistically significant predictors of the likelihood of voicing. These effects are illustrated in figure 6. Voicing tended to occur more often the higher the speech rate ($\beta = -0.43$, $z = -5.85$, $p < 0.0001$), and the shorter the low-band intensity dip ($\beta = -0.71$, $z = -8.3$, $p < 0.0001$). With respect to word position, /s/ tended to be voiced more often in word-final position than in word-initial and medial positions (word-initial: $\beta = -0.46$, $z = -2.87$, $p < 0.005$; word-medial: $\beta = -0.52$, $z = -3.01$, $p < 0.005$). An ANOVA test confirmed that word position conditioned the occurrence of voicing after controlling for the effects of duration ($\chi^2 = 11.06$, $p < 0.005$).

2.2.2 High-Band Intensity Difference

As explained above, we used the maximal difference in intensity between the midpoint of [s] and the beginning of the following vowel in a 4 to 8-kHz band as an indirect measure of the strength of frication in [s]. We first examined the distribution of this high-band intensity difference measure. It exhibited a mean of 17.88 dB, a standard deviation of 8.21 dB and a range extending from -3.75 to 44.2 dB. Around 17% of the tokens had values below 10 dB, and 6% below 5 dB. As an illustration of the incidence of weak frication in our data, consider that the examples of /s/ realizations with little or no weakening in figure 1 have high-band intensity differences of 19.34 and 18.6 dB, which is slightly above average in our data, while the weakened examples in figures 2 and 3 had much lower values ranging from -0.1 to 4.75 dB. These numbers indicate that weak frication such as that seen in figures 2 and 3 is not as frequent as the complete voicing of /s/, which affected over a third of our tokens, although it does occur in a significant number of cases.

We then fitted a regression model by considering all the variables presented in section 2.1.2., including relevant covariates such as speech rate, the segmental context and stress. Since voicing greatly affects the strength of frication, as it diminishes airflow

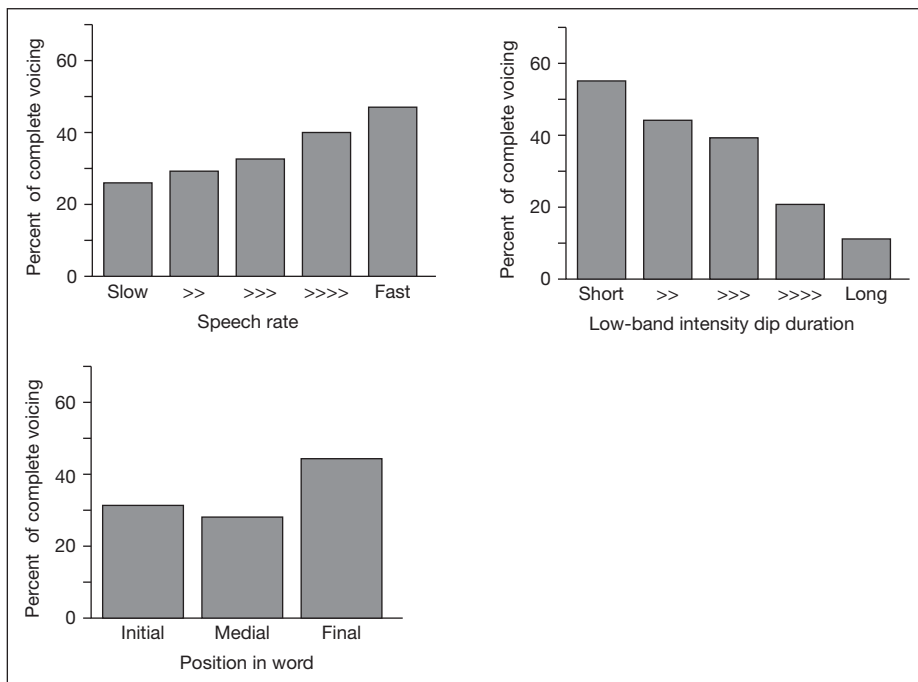


Fig. 6. Percentage of voicing in intervocalic /s/ consonants as a function of speech rate, low-band intensity dip duration (discretized in 5 intervals of equal size), and position in the word.

at the consonantal constriction, we included voicing as a predictor in the regression model. Moreover, we also included low-band dip duration in the model in order to assess if the weakening of frication was associated with the temporal reduction of the consonant.

Speaker was a statistically significant random factor ($\sigma = 3.26$). As for the fixed effects, we found that high-band intensity difference was affected by speech rate, voicing and low-band dip duration. Voiced tokens of /s/ had lower high-band intensity difference values the higher the speech rate ($\beta = -0.65$, $t = -3.53$, $p < 0.0005$). Frication was stronger the longer the duration of the low-band intensity dip, but only for voiceless realizations of /s/ ($\beta = 1.57$, $t = 3.43$, $p < 0.001$). No other statistical effects were observed.

2.2.3 Low-Band Intensity Dip Duration

Finally, we examined the duration of the low-band intensity dip in the vowel-/s/-vowel sequences. Although this measure cannot be used as the segmental duration of /s/, it can be used as an estimate of its temporal reduction (see section 2.1.1 above). The dependent variable exhibited a range from 16 to 134 ms, with a mean of 50 ms and a standard deviation of 21 ms. As an illustration of the amount of temporal reduction of /s/ in our data, consider that around 16% of our tokens had values below 30 ms, which was roughly the durational value of the weakened tokens in figures 2 and 3. As for the

weakening of frication, therefore, temporal reduction of the kind illustrated by figures 2 and 3 appeared to affect a significant number of cases, although it is less common than the complete voicing of /s/.

Regression modeling revealed that speaker and word type were statistically significant random factors, although their effects were rather small ($\sigma = 6$ ms for speaker, and $\sigma = 5$ ms for word). As for the fixed effects, we found that only speech rate, position in the word and stress were statistically significant predictors, with longer low-band intensity dips at slower speech rates ($\beta = 5.44$, $t = 10.15$, $p < 0.0001$), in word-initial and medial positions compared to final position (word-initial: $\beta = 4.82$, $t = 2.75$, $p < 0.01$; word-medial: $\beta = 4.74$, $t = 2.83$, $p < 0.005$), and in stressed syllables ($\beta = 3.55$, $t = 2.52$, $p < 0.05$).

2.3 Discussion

The analyses presented in study 1 show that the voicing of intervocalic /s/ is a frequent phenomenon in our data. Roughly a third of the intervocalic /s/ tokens in our data showed no interruption of voicing. Signs of extreme weakening in terms of loss of frication and temporal compression were also observed, although these phenomena appeared to be somewhat less common than uninterrupted voicing.

We investigated by means of regression modeling which factors influenced the occurrence of /s/ weakening. Number of syllables and the identity of the vowels surrounding /s/ did not yield any effect for any of our dependent variables.

On the other hand, we found that voicing was more frequent at faster speech rates, in word final position, and in shorter consonants. High-band intensity difference, an estimate of the strength of frication, was lower the higher the speech rate, and, within the group of voiceless realizations of /s/, the shorter the consonant. Low-band intensity dip duration, an acoustic estimate of the temporal reduction of intervocalic /s/, was affected by speech rate and position in the word in the same way as voicing, and also showed an effect of stress, with shorter durations in unstressed positions.

Our observations therefore indicate that the weakening of intervocalic /s/ in Spanish is sensitive to several factors known to condition coarticulatory and reduction phenomena, in particular to temporal variables. In our view, this is consistent with the possibility that /s/ weakening arises under the coarticulatory pressure of adjacent vowel segments. The fact, however, that voicing is considerably more common in word-final position suggests that this coarticulatory pattern is not an entirely passive phenomenon, and that it may contain the seeds of phonologization. Our results also indicate that the examined acoustic measures, i.e. the presence of uninterrupted voicing, high-band (4–8 kHz) intensity difference and low-band (0–1 kHz) intensity dip duration, were successful in capturing acoustic variability associated with /s/ weakening.

3 Study 2: Morphological Factors

As explained in the Introduction, Spanish word-final /s/ can function as a plural marker or a verbal suffix. This raises questions regarding possible morphological effects on the realization of /s/ in word-final position, which, interestingly, tends to be more weakened than in other positions in the word (see study 1 above). In study 2, we address two hypotheses regarding the role of morphology on articulation. First,

segments that carry morphosyntactic information by themselves are hypothesized to undergo less reduction than segments that do not. For this reason, /s/ suffixes are predicted to be less prone to weakening than /s/ segments which are part of another morpheme (e.g. *martes* ‘Tuesday’, *es* ‘is’). Second, within the group of suffixes, we hypothesize that weakening should occur more often in /s/ suffixes when they are grammatically or semantically redundant than when they carry new information. For instance, the /s/ plural marker of the word *cosas* ‘things’ is predicted to be more weakened when it is preceded by a word like *las* ‘the’, an article also containing a plural marker, than when it is preceded by a word with no grammatical information regarding number (e.g. *quiero* ‘I want’).

In order to investigate these issues with sufficient statistical power, we created a data set of word-final /s/ consonants consisting of all word-final /s/ consonants in the data set of study 1 plus 572 word-final /s/ consonants additionally sampled from the NCCSp. The additional tokens were measured and annotated in the same way as the tokens from study 1. This data set, exclusively composed of word-final /s/ consonants, contained 930 tokens, of which 559 corresponded to /s/ suffixes (222 plural markers in nouns, 127 plural markers in articles, determiners and pronouns, 59 plural markers in adjectives and 151 verbal suffixes). Of these suffix /s/ tokens, 234 were redundant based on the morphosyntactic or semantic properties of the preceding word (e.g. *las casas* ‘the houses’, *cuatro años* ‘four years’). Most of these redundant suffixes were found in nouns (n = 154) and adjectives (n = 43).

Regression models predicting /s/ voicing, high-band intensity difference and low-band intensity dip duration were fitted with a morphological predictor with three levels (non-suffix /s/, redundant suffix and nonredundant suffix) plus the predictors used in study 1. No differences were observed between any levels of the morphological predictor in the models predicting high-band intensity difference and low-band intensity dip duration. In contrast, the voicing model indicated that nonsuffix /s/ tends to be voiced less often than redundant suffixes (48 vs. 56%; $\beta = 0.39$, $t = 2.01$, $p < 0.05$), and with a similar likelihood as nonredundant suffixes (these were voiced in 50% of the cases; $\beta = 0.01$, $t = 0.06$, $p = 0.94$). A post hoc comparison between redundant and nonredundant suffixes yielded a marginally statistically significant difference ($p = 0.06$). Although the size of the observed differences was relatively small, these results provide some evidence for the hypothesis that grammatical redundancy can lead to increased reduction. On the other hand, the hypothesis that /s/ suffixes are more resistant to weakening than nonsuffix /s/ consonants in general is not supported by our data.

We then investigated to what extent the word position effects on voicing found in study 1 (voicing was more frequent in word-final position) were attributable to morphological redundancy, since many of the word-final consonants in the data set used in this study were morphologically redundant /s/ suffixes. In order to do this, we reran the voicing model of study 1 on a subset of the randomly sampled data set of study 1 from which morphologically redundant /s/ tokens had been excluded. This model showed that, when redundant /s/ consonants are not taken into account, the coefficients of the factor word position remain largely unchanged (word-initial: $\beta = -0.39$, $z = -2.04$, $p < 0.05$; word-medial: $\beta = -0.51$, $z = -2.74$, $p < 0.01$). The coefficients of the other predictors in the model did not change significantly either. Therefore, the frequent voicing of /s/ in word-final position cannot be solely attributed to the fact that it often carries redundant morphosyntactic information, since significant

differences between final and nonfinal /s/ was also observed in a data set from which morphosyntactically redundant /s/ suffixes had been removed.

4 Study 3: Probabilistic and Lexical Factors

We now investigate the potential role of lexical and probabilistic variables on /s/ weakening, including grammatical class (content vs. function words), lexical frequency and contextual predictability, and word-specific effects. Previous research [Bell et al., 2003, 2009; Aylett and Turk, 2004, 2006] has reported effects of lexical and probabilistic factors on global measures of reduction such as word or syllable duration computed across large speech corpora (e.g. higher predictability leads to more reduction). In this study, we assess the extent to which such factors also influence a specific reduction phenomenon such as intervocalic /s/ weakening in spontaneous Spanish.

In order to address this question with sufficient statistical power, we added tokens of highly frequent words to the data set used in study 1. These tokens were measured and annotated in the same way as the tokens from study 1. They were all instances of words which had more than 15 tokens in the randomly sampled data set used in study 1. The extended data set contained 1,615 /s/ consonants, of which more than half ($n = 928$) occurred in 1 of the 12 following high-frequency words: *entonces* ‘then’ ($n = 51$), *es* ‘is’ ($n = 137$), *eso* ‘that’ (pronoun; $n = 96$), *la* (feminine article; $n = 54$), *lo* (masculine article; $n = 65$), *nosotros* ‘we’ ($n = 52$), *sea* ‘to be’ (3rd person singular of the present subjunctive; $n = 57$), *pues* ‘so’ ($n = 52$), *sabes* ‘you know’ ($n = 103$), *se* (pronominal particle; $n = 132$), *si* ‘if’ ($n = 66$) and *sé* ‘I know’ ($n = 50$).

Lexical frequency was computed based on the counts of orthographic forms in the transcriptions of the NCCSp. We assessed the validity of our frequency counts by computing correlations with frequency counts from two other corpora of spoken Spanish, the Spanish Call Home Corpus and the Corpus Oral del Español [Garrett et al., 1997]. In both cases, high correlations were observed ($r = 0.94$ and $r = 0.89$). As a measure of contextual predictability, we computed two word bigram frequencies from the transcriptions of the NCCSp: $C(w_{i-1} w_i)$ (the bigram frequency of the word preceding the target word and of the target word) and $C(w_i w_{i+1})$ (the bigram frequency of the target word and of the word following the target word). All frequency measures were log-transformed, since they exhibited highly skewed distributions. Finally, we considered nouns, adjectives, adverbs and verbs as content words, and all other word types as function words.

The computed probabilistic variables and grammatical class were correlated with each other in several cases. Grammatical class and lexical frequency were moderately associated ($r = 0.33$; function words tended to be of higher lexical frequency), and lexical frequency and the two word bigram frequency measures also displayed moderate correlations ($r = 0.47$ and $r = 0.55$). Since multicollinearity among predictors poses serious obstacles to regression modeling, we decided to run separate analyses for each variable first, and then proceed by fitting pairs of orthogonalized variables using the residuals method (see section 2.2.1 above).

Contrary to our expectations, intervocalic /s/ consonants did not appear to be significantly more prone to voicing in function words than in content words (39.9% for function words vs. 37.3% for content words). This was confirmed by a regression model in which the conditioning factors of voicing identified in section 2.2.1 were

Table 2. High-frequency words: number of tokens analyzed (n), percentage of voiced /s/ and statistically significant p values obtained for the difference between each high-frequency word and low-frequency words in the models predicting voicing, high-band intensity difference and low-band intensity dip duration (see section 2.1.1 for details)

Word	n	Percent voiced	Voicing	Intensity difference	Dip duration
Word-initial and medial position					
<i>es</i> <u>o</u>	96	21			
<i>no</i> <u>s</u> <i>otros</i>	52	57	<0.001	<0.0001	<0.05
<i>se</i> <u>a</u>	57	29		<0.0001	<0.0001
<i>s</i> <u>a</u> <i>bes</i>	52	55	<0.05	<0.05	<0.05
<i>se</i>	132	35		<0.05	<0.05
<i>s</i> <u>i</u>	66	30			
<i>s</i> <u>e</u>	50	32			
Word-final position					
<i>ent</i> <u>on</u> <i>ces</i>	51	49			
<i>e</i> <u>s</u>	137	43			
<i>la</i> <u>s</u>	54	51			
<i>lo</i> <u>s</u>	65	33			
<i>pue</i> <u>s</u>	52	59			
<i>sab</i> <u>e</u> <i>s</i>	51	49			

controlled for ($p > 0.4$). Similarly, no effects of word class were found on high-band intensity difference ($p > 0.35$) or on low-band intensity dip duration ($p > 0.2$). These results suggest that /s/ weakening in spontaneous Spanish is not conditioned by the function versus content word distinction.

We then investigated the roles of lexical frequency and contextual predictability. Again, we started from models including the conditioning factors identified in study 1, to which we added the probabilistic predictors separately, and in orthogonalized pairs. No effect of any of the probabilistic variables was found in any of the analyses predicting voicing ($p > 0.25$ in all cases), nor high-band intensity difference ($p > 0.3$ in all cases). For low-band intensity dip duration, on the other hand, we found a statistically significant effect of word bigram frequency $C(w_i w_{i+1})$, that is of the bigram frequency of the target word and its following word. Contrary to our expectations, intervocalic /s/ consonants tended to have longer low-band intensity dip durations the higher the $C(w_i w_{i+1})$ value ($\beta = 0.6$, $t = 2.2$, $p < 0.05$). It should be noted, however, that the size of this effect was small. As an illustration, consider that the range of $\log C(w_i w_{i+1})$ in our data extended from 0 to 6.3, and that the change in low-band intensity dip duration by $\log C(w_i w_{i+1})$ unit was only of 0.6 ms. In our data, therefore, the maximum predicted difference in low-band intensity dip duration attributable to $C(w_i w_{i+1})$ is only 3.78 ms. Given the small size of this effect, further research would be needed to establish its robustness.

Finally, we investigated word-specific effects. Our goal was to test if specific high-frequency words exhibited more /s/ weakening than predicted by the regression models developed in study 1. We defined a factor *word type* with 13 levels, one for each of the 12 highly frequent words, plus a default level corresponding to all other word types in the corpus. Given that the position of /s/ within the word and word type were not independent, we split the data set into subsets with word-final and non-word-final /s/, and ran separate analyses on each of them. The results are summarized in table 2.

For voicing, we found statistically significant effects for *nosotros* (word-medial: $\beta = 0.92$, $z = 2.68$, $p < 0.001$) and *sabes* (word-initial /s/: $\beta = 0.73$, $z = 2.2$, $p < 0.05$). The word-medial /s/ in *nosotros* and the initial /s/ in *sabes* tended to be voiced more often than /s/ consonants in other words (table 2).

For high-band intensity difference, effects were found again for the word-medial /s/ in *nosotros* ($\beta = -5.25$, $t = -5.05$, $p < 0.0001$), *sea* ($\beta = -3.95$, $t = -4.08$, $p < 0.0001$) and word-initial /s/ in *sabes* ($\beta = -2.23$, $t = -2.12$, $p < 0.05$). High-band intensity difference values tended to be lower for these words than for the group of low-frequency words.

As for low-band intensity dip duration, effects were observed for the words *nosotros* (word-medial: $\beta = -6.57$, $t = -2.33$, $p < 0.05$), *sea* ($\beta = -14.82$, $t = -5.61$, $p < 0.0001$), *se* ($\beta = -5.23$, $t = 2.27$, $p < 0.05$), and for the word-initial /s/ in *sabes* ($\beta = -6.51$, $t = -2.25$, $p < 0.05$). Low-band intensity dip duration tended to be shorter in these words than in low-frequency words.

In conclusion, our data do not seem to support the hypothesis that /s/ consonants in highly frequent and predictable words and function words generally exhibit more weakening than when they occur in non-high-frequency words and content words. However, word-specific trends towards more reduction could be found in some high-frequency words, in particular in the words *nosotros* and *sabes*.

5 General Discussion

The studies presented above have investigated the weakening of intervocalic /s/ consonants in the NCCSp, a newly created corpus of informal conversations among Spanish speakers from Madrid. Since acoustic segmentation following traditional procedures proved inappropriate for our data, we devised three acoustic parameters related to the articulation of /s/ that could be measured semiautomatically in a relatively large corpus of spontaneous speech such as the NCCSp. The analysis of such parameters (high-band intensity difference, low-band intensity dip duration and the occurrence of uninterrupted voicing) have allowed us to assess the pervasiveness of /s/ weakening in this variety of Spanish, and the relevance of some of its conditioning factors.

In study 1, we have observed that over a third of the /s/ consonants randomly sampled from the NCCSp exhibit uninterrupted voicing throughout their realization, and that the frequency of this phenomenon is highly variable among speakers. Judging from the distributions of high-band intensity difference and low-band intensity dip duration values, it appears that extreme reduction in terms of frication and the durational properties of intervocalic /s/ is less common than the occurrence of uninterrupted voicing. This suggests that /s/ weakening may be primarily an issue of glottal coarticulation between intervocalic /s/ and surrounding vowels, and that the loss of frication may be largely or entirely driven by the presence of intrusive voicing throughout the fricative consonant [Ohala and Solé, 2010]. However, since cases of temporal reduction, which probably involved reduced supraglottal gestures, were also present in our data, we cannot discard the possibility that /s/-weakening has independent glottal and supraglottal components. Articulatory and aerodynamic data would be needed to properly address this issue.

In study 1 we have also examined to what extent /s/ weakening is affected by several variables known to affect coarticulatory and reduction phenomena. All three acoustic parameters were consistently affected by speech rate. As expected, we

observed more weakening at higher speech rates. In order to further investigate the extent to which the voicing and loss of frication in intervocalic /s/ is sensitive to temporal reduction, we also examined the relationship between these dependent variables and low-band intensity dip duration while controlling for speech rate. In both cases, more weakening was observed in temporally reduced consonants. In our view, this is consistent with a view of /s/ weakening as a coarticulatory phenomenon.

Voicing and low-band intensity dip duration were also affected by position in the word, with more weakening in word-final position than in word-initial and medial positions. Although the sort of /s/-weakening observed in our Madrid Spanish data and the /s/ aspiration prevalent in other dialects are not the same phenomena – in our data all /s/ tokens were clearly produced as sibilants – it is interesting to note that the aspiration of intervocalic /s/ in /s/-aspirating dialects occurs almost exclusively in word-final position. Word-final position therefore appears as a recurrent locus of different kinds of /s/ reduction across different Spanish dialects. The fact that /s/ weakening in our Madrid Spanish data is significantly more common in word-final position than in other word positions further suggests that this phenomenon is not entirely passive, and that it may already contain the seeds of a possible sound change analogous to that undergone by /s/-aspirating Spanish dialects.

Stress had a small effect on low-band intensity dip duration. We found that /s/ had slightly longer durations in stressed syllables than in unstressed syllables. Interestingly, no effects of stress were observed for voicing and high-band intensity difference. Although this finding may seem counterintuitive, it parallels our previous observations for intervocalic /p/, /t/ and /k/ consonants in the NCCSp [Torreira and Ernestus, 2011]. The realization of these consonants in spontaneous Spanish showed an effect of stress on duration and constriction degree, but, as in the case of intervocalic /s/, not on the presence of intrusive voicing from adjacent vowels. It therefore appears that stress does not play an important role in the occurrence of intervocalic voicing in voiceless obstruents in Spanish, although it does affect other aspects of their articulation.

Together with previous data on the weakening of Spanish intervocalic voiceless stops [Machuca-Ayuso, 1997; Lewis, 2001; Martínez Celdrán, 2009; Torreira and Ernestus, 2011; Hualde et al., 2011], our findings suggest that the kind of /s/ weakening studied here is probably an instance of a more general lenition process affecting all Spanish intervocalic voiceless obstruents. We now have evidence that several Spanish voiceless obstruents (/p/, /t/, /k/ and /s/) are often realized with uninterrupted voicing when occurring in intervocalic position. Moreover, we have evidence that voiceless stops are often produced without complete oral closures, and that /s/ consonants may exhibit very weak frication noise. In addition, preliminary data from an ongoing study indicates that other Spanish voiceless fricative types in the NCCSp are also subject to frequent intervocalic voicing (25.9% for /f/, 28.7% for /θ/ and 15.5% for /x/). It is possible that all of these cases of consonantal weakening share a common set of causes. For instance, it is possible that the glottal and supraglottal targets of voiceless obstruents in Spanish have all shifted towards more lenited positions in phonetic space. This may be due to global changes in the basis-of-articulation of this language, for instance towards a more sonorant overall quality (e.g. a preponderance of wider constrictions and voicing). Another possibility is that it is a change in the resistance of Spanish voiceless obstruents to the coarticulatory influence from adjacent vowels, not a change in their intrinsic phonetic targets, which is mainly responsible for their weakening. In this case, realizations of Spanish intervocalic /s/ in contexts that

minimize coarticulation should attain supraglottal and glottal targets comparable to those of intervocalic /s/ in languages that do not exhibit significant weakening under conditions more prone to coarticulation (e.g. fast and spontaneous speech as opposed to slow and read speech). A combination of corpus, experimental and quantitative modeling research work would be required to investigate these issues.

One of the goals of the present work was to examine the potential influence of morphological factors on /s/ weakening. We hypothesized that /s/ segments which are part of a larger morpheme (e.g. *martes* ‘Tuesday’) are more prone to weakening than /s/ segments that carry grammatical information by themselves as a plural or verbal suffix. We also hypothesized that /s/ suffixes that are redundant based on the linguistic context (e.g. *las casas* ‘the houses’) are more prone to weakening than other word-final /s/ segments. We found some support only for the second hypothesis, with 56% of voicing for redundant /s/ suffixes versus 48% for nonsuffix /s/ and 50% for nonredundant suffixes. Our findings are therefore in line with Poplack [1980], Hundley [1987] and Erker [2010], who found that the reduction of /s/ in Puerto Rican, Peruvian and Dominican Spanish is not conditioned directly by morphological factors, but rather by the contextual information available in the utterance.

Regarding the roles of probabilistic and lexical factors, our data do not support the hypothesis that highly frequent and predictable words and function words systematically exhibit more reduction than low-frequency, unpredictable words and content words. In order to test for lexical frequency and word class effects, we used an extended data set containing numerous tokens of highly frequent lexical items, many of them function words. In spite of the increase in statistical power allowed by this data set, lexical frequency and word class did not approach statistical significance in the analyses of voicing and high-band intensity difference. In the case of low-band dip intensity duration, our estimate of temporal reduction, we found a small effect of predictability based on the following word [word bigram frequency $C(w_i, w_{i+1})$], but this effect, as for two different acoustic correlates of /s/ in Caleño Spanish [File-Muriel and Brown, 2011], was in the opposite direction of that predicted by the probabilistic reduction hypothesis. Given the small size of this effect, further research is required to establish whether it is recurrent across different data sets.

One possible reason why our findings on the role of lexical class, frequency and contextual predictability contrast with much of the previous literature is that our data set did not contain a large number of word tokens and word types. In this sense, for instance, our data set is not comparable to those used in influential studies like those of Bell et al. [2009] and Aylett and Turk [2004]. A second possible explanation is that we focused on three acoustic parameters of a very specific reduction phenomenon, rather than on global measures of reduction such as word or syllable duration. This would suggest that the frequency and word class factors claimed to favor reduction in previous research may well exhibit general effects on global phonetic parameters measured across large-scale corpora, but not necessarily on any reduction phenomena when considered individually. Under the latter hypothesis, the reasons why some reduction phenomena are more sensitive to lexical and probabilistic factors than others would need to be investigated. In our case, intervocalic /s/ weakening appeared to be very frequent in our data of conversational Spanish. It is possible that under such circumstances ceiling effects inhibit lexical class and probabilistic factors from affecting articulation.

A third possibility is that the general grammatical class and probabilistic effects reported by previous research are largely carried by a number of specific words and

word sequences. This would be consistent with our findings of word-specific effects for several lexical items, particularly *nosotros* ‘we’ and *sabes* ‘you know’. These words exhibited more /s/ weakening (more frequent voicing, lower high-band intensity differences, and shorter low-band intensity dip durations) than the bulk of low-frequency words in the data. Since *nosotros* is a very long high-frequency function word, it could be that frequency or word class do condition /s/ weakening, but only in interaction with word length. As for *sabes*, its status as a very frequent discourse marker may explain its special behavior. In our view, the fact that a number of word-specific effects could be identified, in contrast to general effects of lexical class and frequency, suggests that the influence of specific lexical properties (e.g. grammatical class, frequency, contextual predictability) on articulation is more complex than is usually acknowledged, and that further investigation on the role of these factors in articulation is still needed.

In conclusion, this article has provided a description of a reduction phenomenon little studied in the past despite its pervasiveness in conversational Spanish. Statistical analysis of three acoustic correlates of intervocalic /s/ weakening has provided insights into the role of several of its conditioning variables. The weakening of /s/ was affected by temporal and prosodic factors, suggesting that it may be caused by coarticulation with adjacent vowels. We did not find that /s/ suffixes are more resistant to weakening than word-final /s/ consonants in monomorphemic words. In fact, we found that /s/ suffixes tend to be slightly more weakened than other word-final /s/ tokens when they occur in morphosyntactically redundant contexts. Regarding the roles of lexical class, frequency and predictability, often reported in the literature as important conditioning factors of speech reduction, we have observed a more complex pattern of variability than initially expected. Further phonetic research of the kind presented here is needed to better understand how all these factors come to condition speech reduction.

Acknowledgements

We would like to thank Louis ten Bosch for his assistance with HTK, and Jessamyn Schertz, Miquel Simonet, Daniel Recasens and three anonymous reviewers for their comments on manuscript versions of this article. This work was supported by a European Young Investigator Award to the second author.

References

- Adda-Decker, M.; de Mareüil, P.B.; Adda, G.; Lamel, L.: Investigating syllabic structures and their variation in spontaneous French. *Speech Commun.* 46: 119–139 (2005).
- Adda-Decker, M.; Snoeren, N.D.: Quantifying temporal speech reduction in French using forced speech alignment. *J. Phonetics* 39: 261–270 (2011).
- Aylett, M.; Turk, A.: The smooth signal redundancy hypothesis: A functional explanation for relationships between redundancy, prosodic prominence, and duration in spontaneous speech. *Lang. Speech* 47: 31–56 (2004).
- Aylett, M.; Turk, A.: Language redundancy predicts syllabic duration and the spectral characteristics of vocalic syllable nuclei. *J. acoust. Soc. Am.* 119: 3048 (2006).
- Baayen, H.: *LanguageR: Data sets and functions with analyzing linguistic data: A practical introduction to statistics.* R package version 0.953 (2008).
- Baker, R.E.; Baese-Berk, M.; Bonnasse-Gahot, L.; Kim, M.; Engen, K.J.V.; Bradlow, A.R.: Word durations in non-native English. *J. Phonetics* 39: 1–17 (2011).
- Bates, D.; Sarkar, D.: *lme4: Linear mixed-effects models using Eigen and R syntax.* R package version 0.9975-10 (2006).

- Bell, A.; Brenier, J.; Gregory, M.; Girand, C.; Jurafsky, D.: Predictability effects on durations of content and function words in conversational English. *J. Mem. Lang.* 60: 92–111 (2009).
- Bell, A.; Jurafsky, D.; Fosler-Lussier, E.; Girand, C.; Gregory, M.; Gildea, D.: Effects of disfluencies, predictability, and utterance position on word form variation in English conversation. *J. acoust. Soc. Am.* 113: 1001–1024 (2003).
- Brown, E.K.: The relative importance of lexical frequency in syllable- and word-final /s/ reduction in Cali, Colombia; in Collantine, Selected proceedings of the 11th Hispanic Linguistics Symposium, pp. 165–178 (Cascadilla Proceedings Project, Somerville 2009).
- Brown, E.L.; Torres Cacoullos, R.: Spanish /s/: A different story from beginning (initial) to end (final); in Núñez-Cedeño, López, Cameron, A Romance perspective in language knowledge and use. Selected papers from the 31st linguistic symposium of Romance languages (LSRL), pp. 22–38 (John Benjamins, Somerville 2003).
- Bürki, A.; Fougeron, C.; Gendrot, C.; Frauenfelder, U.H.: Phonetic reduction versus phonological deletion of French schwa: Some methodological issues. *J. Phonetics* 39: 279–288 (2011).
- Bybee, J.: Phonology and language use (Cambridge University Press, Cambridge 2001).
- Bybee, J.: Word frequency and context of use in the lexical diffusion of phonetically conditioned sound change. *Lang. Var. Change* 14: 261–290 (2002).
- Erker, D.G.: A subsegmental approach to coda /s/ weakening in Dominican Spanish. *Int. J. Sociol. Lang.* 203: 9–26 (2010).
- Ernestus, M.: Voice assimilation and segment reduction in Dutch: A corpus-based study of the phonology-phonetics interface (LOT, Utrecht 2000).
- Ernestus, M.; Lahey, M.; Verhees, F.; Baayen, H.: Lexical frequency and voice assimilation. *J. acoust. Soc. Am.* 120: 1040–1051 (2006).
- File-Muriel, R.: A study of lenition: The role of lexical frequency and phonetic context in the weakening of lexical /s/ in the Spanish of Barranquilla; PhD thesis Indiana University, Bloomington (2007).
- File-Muriel, R.J.; Brown, E.K.: The gradient nature of s-lenition in Caleño Spanish. *Lang. Var. Change* 23: 223–243 (2011).
- Garrett, S.; Morton, T.; McLemore, C.: LDC Spanish lexicon (Linguistic Data Consortium, University of Pennsylvania, Philadelphia 1997).
- Guy, G.R.: Contextual conditioning in variable lexical phonology. *Lang. Var. Change* 3: 223–239 (1994).
- Hernández-Campoy, J.M.; Trudgill, P.: Functional compensation and Southern Peninsular Spanish /s/ loss. *Folia ling. hist.* 23: 31–57 (2002).
- Hoole, P.: Laryngeal coarticulation: Section A; in Hardcastle, Hewlett, Coarticulation: Theory, data and techniques, pp. 105–121 (Cambridge University Press, Cambridge 1999).
- Hualde, J.I.: Unstressed words in Spanish. *Lang. Sci.* 31: 199–212 (2009).
- Hualde, J.I.; Simonet, M.; Nadeu, M.: Consonant lenition and phonological recategorization. *J. Lab. Phonol.* 2: 301–329 (2011).
- Hundley, J.E.: Constraints on plural marker deletion in Peruvian Spanish. *Hispania* 70: 891–894 (1987).
- Johnson, K.: Massive reduction in conversational American English; in Yoneyama, Maekawa, Spontaneous speech: Data and analysis. Proceedings of the 1st Session of the 10th International Symposium, pp. 29–54 (The National International Institute for Japanese Language, Tokyo 2004).
- Kohler, K.: Investigating unscripted speech: Implications for phonetics and phonology. *Phonetica* 57: 85–94 (2000).
- Kuperman, V.; Pluymaekers, M.; Ernestus, M.; Baayen, H.: Morphological predictability and acoustic duration of interfixes in Dutch compounds. *J. acoust. Soc. Am.* 121: 2261–2271 (2007).
- Labov, W.: Quantitative analysis of linguistic variation; in Ammon, Dittmer, Mattheier, Trudgill, Sociolinguistics: An international handbook of the science of language and society, pp. 6–21 (De Gruyter, Berlin 2004).
- Lafford, B.: Is functionalism a fact? Data from the Caribbean. *Hisp. Linguistics* 3: 49–74 (1989).
- Lewis, A.: Weakening of intervocalic /p, t, k/ in two Spanish dialects: Toward the quantification of lenition processes; PhD thesis, University of Illinois at Urbana-Champaign (2001).
- Machuca Ayuso, M.J.: Las obstruyentes no continuas del español: relación entre las categorías fonéticas y fonológicas en habla espontánea; PhD thesis, Universitat Autònoma de Barcelona (1997).
- Martínez Celdrán, E.: Sonorización de las oclusivas sordas en una hablante murciana: problemas que plantea. *Est. Fonét. Exp.* 18: 255–271 (2009).
- Minnick Fox, M.: Usage-based effects in Latin American Spanish syllable-final /s/ lenition; PhD thesis, University of Pennsylvania, Philadelphia (2006).
- Moreno Sandoval, A.; Toledano D.T.; de la Torre, R.; Garrote, M.; Guirao, J.M.: Developing a phonemic and syllabic frequency inventory for spontaneous spoken Castilian Spanish and their comparison to text-based inventories; in Calzolari, Choukri, Maegaard, Mariani, Odjik, Piperidis, Tapias, Spontaneous speech: Data and analysis. Proceedings of the Sixth International Language Resources and Evaluation (LREC'08), pp. 1097–1100 (European Language Resources Association, Marrakech 2008).
- Ohala, J.J.; Solé, M.J.: Turbulence and phonology; in Fuch, Toda, Zygis, Turbulent sounds: An interdisciplinary guide, pp. 37–101 (De Gruyter Mouton, Berlin 2010).
- Pluymaekers, M.; Ernestus, M.; Baayen, H.: Articulatory planning is continuous and sensitive to informational redundancy. *Phonetica* 62: 146–159 (2005a).
- Pluymaekers, M.; Ernestus, M.; Baayen, H.: Lexical frequency and acoustic reduction in spoken Dutch. *J. acoust. Soc. Am.* 118: 2561–2569 (2005b).

- Poplack, S.: Deletion and disambiguation in Puerto Rican Spanish. *Language* 56: 371–385 (1980).
- Schuppler, B.; Ernestus, M.; Scharenborg, O.; Boves, L.: Acoustic reduction in conversational Dutch: A quantitative analysis based on automatically generated segmental transcriptions. *J. Phonetics* 39: 96–109 (2011).
- Shadle, C.H.; Scully, C.: An articulatory-acoustic-aerodynamic analysis of [s] in VCV sequences. *J. Phonetics* 23: 53–66 (1995).
- Son, R. van; Pols, L.: An acoustic model of communicative efficiency in consonants and vowels taking into account context distinctiveness. *Proceedings of the 15th International Congress of Phonetic Sciences, Barcelona*, pp. 2141–2144 (2003).
- Stevens, K.N.: *Acoustic phonetics* (The MIT Press, Cambridge 1998).
- Torreblanca, M.: La ‘s’ sonora prevocálica en el español moderno. *Thesaurus* 41: 59–69 (1986).
- Torreira, F.; Adda-Decker, M.; Ernestus, M.: The Nijmegen Corpus of Casual French. *Speech Commun.* 52: 201–221 (2010).
- Torreira, F.; Ernestus, M.: The Nijmegen Corpus of Casual Spanish; in Calzolari, Choukri, Maegaard, Mariani, Odijk, Piperidis, Rosner, Tapias, *Proceedings of the Seventh Conference on International Language Resources and Evaluation (LREC’10)* (European Language Resources Association, Valletta 2010a).
- Torreira, F.; Ernestus, M.: Phrase-medial vowel devoicing in spontaneous French. *Proceedings of the 11th Annual Conference of the International Speech Communication Association (Interspeech 2010)*, Makuhari, Japan, pp. 2006–2009 (2010b).
- Torreira, F.; Ernestus, M.: Realization of voiceless stops and vowels in conversational French and Spanish. *J. Lab. Phonol.* 2: 331–353 (2011).
- Young, S.; Evermann, G.; Hain, T.; Kershaw, D.; Moore, G.; Odell, J.; Ollason, D.; Povey, D.; Valtchev, V.; Woodland, P.: *The HTK Book 3.2* (Entropic, Cambridge 2002).