1. Introduction

One of the phonological processes in Spanish is the regressive voicing assimilation of /s/ when it is followed by a voiced consonant (Quilis 1999; Navarro-Tomás 2004; Hualde 2005; Schwegler, Kempff & Ameal-Guerra 2010). Thus, the segment /s/ is produced as [s] in front of a voiceless consonant, as in *disco* ([disko], ‘disc’), while it can be produced as [z] in front of a voiced consonant, as in *isla* ([izla], ‘island’) or *mismo* ([mizmo], ‘same’). This process is not categorical, although Schwegler et al. (2010) claim that the application of the rule is the most frequent option. Hualde (2005) also says that "this process is optional and it is also often incomplete; that is, the fricative may become only partially voiced" (107).

However, although the aforementioned authors indicate that the rule is not categorical and that the assimilation process does not occur in all cases, they do not mention the reasons why this occurs or whether there are specific factors that determine whether there is voicing assimilation or not.

In this study, we want to explore if this assimilation rule is affected by different linguistic and extralinguistic factors, that is, if the variation between the application and non-application of the assimilation process is due to different factors and, in that case, which factors. In order to achieve this, we analyzed the production of /s/ before voiced consonants, both word internally and across word boundaries, in a bilingual region (Basque Country). The data used for the study were collected in the regional capital, Vitoria, a city of around 240,000 inhabitants. The main goal of this study is to analyze whether the rate of voicing assimilation is the same before all consonants or whether there are some consonants that result in higher rates of voicing assimilation than others. The fact that the data were collected in the Basque Country using the exact same methodology and instrument than that used by Schmidt & Willis (2011) for Mexican Spanish allows for a cross-dialectal comparison between this variety and Basque Spanish, since the variety of methodologies used in previous studies makes this comparison difficult.

This paper starts with a review of the phonological processes in Spanish and Basque and a review of previous literature on the topic. Following those sections, we present the methodology and the results, as well as a discussion on the findings. Finally, we include a conclusion where we present a summary of the findings, as well as comments on future directions for the study.
2. Sibilants and voicing assimilation in Spanish and Basque

2.1. Voicing assimilation in Spanish

In the voicing assimilation process, the voicing of the following consonant is transferred to the /s/ segment causing it to become voiced. The assimilation rule proposed by Schwegler et al. (2010: 171) is, thus, the following:

(1)     Orthographic "s"    
[ z ] /  _ _  C [+son]  
Orthographic "z"

Following this rule, the /s/ in a word like revistas (‘magazines’) would be voiceless, since it is followed by a voiceless segment (/t/) (Figure 1). On the other hand, the /s/ in a word like bisnietos would become voiced (that is, the assimilation rule would apply), since the segment is followed by a voiced consonant /n/ (Figure 2). There are several acoustic correlates connected with voicing: the presence of a voice bar at the bottom of the spectrogram, the presence of periodicity in the waveform and the presence of the fundamental frequency (F₀).

Figure 1. Voiceless /s/ preceding a voiceless consonant /t/ in the word revistas (Participant S1).

As can be seen in Figure 1, the /s/ is voiceless for the most part, since it precedes a voiceless consonant [t]. However, some voicing can be observed at the beginning of the segment, due a coarticulatory process with the previous vowel. This process was also observed by Schmidt & Willis (2011) in their data from Mexico.

In Figure 2 the segment, followed by a voiced consonant [n], is completely voiced, indicated by the presence of an uninterrupted voice bar and the periodicity in the waveform.
Figure 2. Voicing assimilation of /s/ in front of a voiced consonant /n/ in the word *bisnietos* (‘great-grandchildren’) (Participant S1).

This assimilation does not occur word-internally only. It can also appear across word boundaries or, as Schwegler et al. (2010) state, "a lo largo de todo el grupo fónico" (‘throughout the whole phonic group’, my translation) (172). An example suggested by the authors is *muchas gracias* ([‘mu-t/g1877az-’/g1845/g1872a-/g86/g76/g2156/g68/g86/g64, ‘thank you very much’). It is also important to notice that this process is usually associated with consonants. However, the process can also occur in intervocalic position, as Torreblanca (1978), Schmidt & Willis (2011) and Torreira and Ernestus (2012) found, and in the previous literature on the topic, to the author’s knowledge, only Muñiz Cachón and Cuevas Alonso (2012) have investigated the role of the following consonant.

Several studies have already analyzed the voicing assimilation process. Torreblanca (1978), for example, used a sonograph to analyze /s/ voicing in Toledo with spontaneous speech. In his data Torreblanca found voicing both before voiced consonants and in intervocalic position, although more voicing was found before voiced consonants than in intervocally. However, despite the lower frequency of /s/ voicing in intervocalic position when compared to preconsonantal position, Torreblanca (1986) argues that it is a widespread phenomenon across many Spanish speaking regions both in Spain and in Latin America. Additionally, he explains that, despite what other authors have suggested, it is not a case of phonetic archaism, but a modern innovation in the language (except for a handful of exceptions). Torreblanca (1978, 1983) also found that the /s/ could be either partially or completely voiced and argued that /s/ voicing is due to articulatory relaxation, connected with spontaneous, relaxed and fast speech. Schmidt & Willis (2011) analyzed voicing assimilation in Mexico. For their study they created the instrument that we adapted and used in the current study. They recorded 12 university students of the Universidad Autónoma Metropolitana-Iztapalapa (UAM-I) (6 males and 6 females). In the study they found that the voicing assimilation process is not categorical, to the point that in 37% of the expected contexts there was no voicing assimilation, as well as instances of s-voicing in unexpected contexts, such as intervocalic position, where 9% of the /s/ segments were voiced. As previously mentioned, the general literature about the process only mentions the process as occurring in preconsonantal position, hence the idea of the intervocalic position being "unexpected". They also found that some of the voicing of the previous vowel transferred to the
beginning of the voiceless sibilant segment due to a coarticularatory process and that there is a tendency (not statistically significant) for more assimilation in the case of male speakers. Finally, they showed that there was an effect of phrase position in the phonological process.

Campos-Astorkiza (2010) analyzed voicing assimilation in a Northern Peninsular variety of Spanish. For her study, the participants (6 female speakers) had to read a list of stimuli in a carrier sentence *Dice que _____ todo* (‘he/she says that _____ everything’). In all the tokens the /s/ was followed by a voiced or voiceless stop. Campos-Astorkiza found that stress did not have an effect on the degree of voicing (although it did have an effect on the duration of the frication) and that manner of articulation had an effect. Open approximants (where there is a closure but without air compression, according to Campos-Astorkiza 2010, citing Martínez-Celdrán 1991, 1998) resulted in a higher rate of voiced /s/.

Sedó (2012) analyzed the data from the same speakers in the current study to determine if there were differences in the application of the assimilation process depending on the position of the /s/ in the word (assimilation word internally vs. across word boundaries) and on stress. Sedó found that there was a higher assimilation rate across word boundaries than word internally (49% vs. 32% respectively) and that the assimilation rate was higher when the /s/ segment was in an unstressed syllable, although the difference with respect to stressed syllables was minimal (46% vs. 40%). However, Sedó (2012) did not perform any statistical analysis and considered voicing as a binary variable rather than as a continuous variable.

Torreira and Ernestus (2012) used a random sample from the Nijmegen Corpus of Casual Spanish to analyze the weakening of the intervocalic /s/. Torreira and Ernestus (2012) carried out three different studies to investigate the incidence of /s/ weakening and its conditioning factors (speech rate, position of /s/, stress, segmental contact and comparison with other fricatives) and the potential role of morphosyntactic and lexical factors as well as frequency effects. One third of the consonants showed uninterrupted voicing. There were higher rates with faster speech, word finally, and when the following (but not adjacent) consonant was a voiceless fricative. No general effects of word frequency or word class (content vs. function words) were found, although there were interactions between word frequency and word class (function words showing higher frequencies than content words).

Muñiz Cachón and Cuevas Alonso (2012) analyzed the process in Asturian Spanish with data from a reading task. They categorized the realizations as voiced, voiceless and mixed (according to the authors, for those cases when it was not clear whether it was voiced or voiceless, although they did not give any other explanation regarding this categorization). The authors found that 51.4% of the tokens were voiced and that males presented higher voicing rates than females, which the authors attributed to females producing more tense and less lazy realizations, which would be less subject to becoming voiced. The results from Muñiz Cachón and Cuevas Alonso’s study also showed differences according to the consonants following the /s/ and found that there was more voicing when the /s/ was followed by [j] and [β], less voicing when it was followed by [ð] [ɣ] and [l] and even less when it was followed by [m] and [n].

The assimilation process in Ecuador has also been the object of study (Lipski 1989, Bradley 2005, Colina 2009, Chappell 2011). Most of these authors offer a description of the process in Ecuador, in some cases using Dispersion Theory (Bradley 2005) and in others from the perspective of the Optimality Theory (Colina 2009). Chappell (2011), however, did use empirical data to analyze voicing in Ecuador. Chappell focused on intervocalic voicing and found that different factors determined the rate of /s/ voicing in intervocalic position depending on the position of the /s/ (word medial, word final and word initial). Across word boundaries voicing almost categorically occurred when the /s/ appeared word finally (V_#V) (91%) as opposed to word initially (V/#V) (6%). Word internally the voiceless variant was favored as well (89%). Additionally, the following vowel and word frequency were influential factors in word-medial tokens (frequent words and vowels /o/ and /a/ favored the voiced variante). It is unclear, however, how word frequency was measured and categorized. This article provided clear empirical data on intervocalic voicing in this variety of Spanish. García (2013) investigated if intervocalic voicing also occurred in Highland Colombian Spanish, arguing similarities between Highland Colombian Spanish and Highland Ecuadorian Spanish. However, she found that, although it does occur, its frequency is considerably lower than in Ecuador and is correlated with speech rate, as is /s/ voicing in preconsonantal position, as well.
To sum up, as Schwegler et al. (2010) mention, the application of the assimilation process is expected to be the most frequent outcome and, according to Hualde (1995), the process is optional and there can be only partial voicing. But, although these two authors only mention voiced consonants as the segments that determine the voicing of the /s/ segment, evidence from Schmidt & Willis (2011) and Torreira and Ernestus (2012) indicate that vowels also affect it and results from Schmidt & Willis (2011), Torreira and Ernestus (2012) and Sedó (2012) suggest that the process is variable depending on factors such as gender, position in the word and speech rate. It remains to be confirmed, however, if the phonetic context (previous vowel and following consonant) also has an effect on the application of the assimilation process.

2.2. Voicing assimilation in Basque

According to previous literature, voicing processes in Basque are variable. On the one hand it has been argued that sibilant voicing assimilation occurs before a voiced consonant, as opposed to other contexts (before voiceless consonants, before pauses, and before vowels across word boundaries due to the resyllabification) (Hualde & Bilbao, 1992, Hualde & Ortiz de Urbina, 2003). On the other hand, voiced stops are devoiced after a voiceless segment, although whether it affects all three voiced stops or not will depend on the dialect (Hualde, 1991). When a sibilant is followed by a voiced stop, both sibilant voicing and stop devoicing can be found (Hualde & Ortiz, 2003). The direction of the effect will depend on the context. Included in Table 1 are some of the contexts where the sequence sibilant + voiced stop can be found, as well as the outcomes for each context. It is important noticing that the sequence is not found morpheme internally (Hualde & Ortiz, 2003).

Table 1. Examples of outcomes of the sequence sibilant + voiced stop (taken from Hualde & Ortiz de Urbina, 2003).

<table>
<thead>
<tr>
<th>Context</th>
<th>Example</th>
<th>Realization</th>
<th>Voicing Process</th>
</tr>
</thead>
<tbody>
<tr>
<td>Suffixation</td>
<td>/aots/ + /gaba/</td>
<td>[aoaɾkazbe] ‘voiceless’</td>
<td>Stop devoicing</td>
</tr>
<tr>
<td></td>
<td>‘voice’ + ‘without’</td>
<td></td>
<td></td>
</tr>
<tr>
<td>After the prefix des-</td>
<td>/dežberđiňa/</td>
<td>[deʒβerđiňa]</td>
<td>Sibilant voicing</td>
</tr>
<tr>
<td></td>
<td>‘different’</td>
<td></td>
<td></td>
</tr>
<tr>
<td>Across word boundaries</td>
<td>/eškuž garbitu/</td>
<td>[eškužyaɾbitu]</td>
<td>Sibilant voicing</td>
</tr>
<tr>
<td></td>
<td>‘hand washed’</td>
<td>[eškušyaɾbitu]</td>
<td></td>
</tr>
<tr>
<td>Ez + conjugated verb</td>
<td>/ešbaดา/ ‘if it isn’t’</td>
<td>[ešpaดา]</td>
<td>Stop devoicing</td>
</tr>
<tr>
<td></td>
<td></td>
<td>[ežbaดา]</td>
<td>Sibilant voicing</td>
</tr>
</tbody>
</table>

Although Vitoria is mainly Spanish-dominant, knowledge of Basque is institutionalized in the city, as is in the rest of the Basque Country. In schools Basque is either the instruction language or a mandatory subject, and public servants are required to demonstrate different levels of Basque depending on their job. Additionally, the number of active bilingual speakers has increased over the last 20 years (Viceconsejería de política lingüística del Gobierno Vasco, 1997, 2008, 2011) and the number of students registered for bilingual instruction or instruction almost entirely in Basque has also increased (Diputación Foral de Álava, 2009).

The increase in the degree of exposure to Basque by the population of Vitoria and, therefore, of contact between both languages, has the potential to create patterns in Vitoria Spanish that differ from those found in monolingual varieties. In the case of voicing assimilation, it could possible that the variety under study shows differences with respect to surrounding monolingual varieties of Spanish, although in order to confirm it research needs to be done to determine the degree of voicing

1 The term ‘sibilant’ in this section refers only to the two alveolar sibilants found in Basque (/ʃ/ and /ʒ/) and not to the postalveolar sibilant (/ʃʃ/).
2 Students choose between three different schooling models: A (all subjects in Spanish but one, which is Basque language), B (bilingual education) and D (all subjects in Basque but one, which is in Spanish).
assimilation in Basque using empirical data, as well as whether different levels of exposure to Basque and knowledge of the language lead to differences in voicing rates in Spanish.

3. Current study

The main objective of this study was to investigate the variability in the voicing assimilation process, by analyzing whether the /s/-voicing assimilation rates change depending on the following voiced consonant or the previous vowel, as well as to investigate which other factors play a role in /s/ voicing.

Based on these goals we have the following research questions:
1) Does the following consonant have an effect on the voicing assimilation rate?
2) Does the preceding vowel have an effect on the voicing assimilation rate?
3) Does the rate of assimilation differ from the rate found in Mexico by Schmidt and Willis (2011)?
4) Which other factors play a role in the voicing assimilation process?

For these research questions we have the following hypotheses:
1) Since Muñiz Cachón and Cuevas Alonso (2012) found differences in the voicing rate depending on the following consonant (more voicing before [ʃ], [β], [θ] and [γ] and less before [l], [m] and [n]), similar differences are expected to be found in the present study.
2) Since there is no previous literature that explains whether there is variation in the assimilation process depending on the previous vowel, we assume that the assimilation rates are going to be similar after the different vowels.
3) The assimilation rate might be similar to that found in Mexico by Schmidt & Willis (2011), since there is little information on the assimilation process in Basque and it is still early to determine whether there will be an influence of Basque. Moreover, the majority of the linguistic consequences of language contact between Basque and Spanish have been of morphosyntactic nature.
4) As for the other factors included in the analysis, some differences are expected to be found with respect to gender, given the differences found between males and females by Schmidt and Willis (2011) and Muñiz Cachón and Cuevas Alonso (2012), as well as depending on the position of the word. In this respect, it is expected to find more voicing word finally, since it is a more productive position in regards to weakening processes (consonant deletion and vowel devoicing, for example). As for stress, given that Campos-Astorkiza (2010) did not find differences and the preliminary results from Sedó (2012) indicate small differences in the voicing rate between stressed and unstressed syllables, no differences are expected in the present study either.

In order to answer these research questions two different analyses were performed. The first, considering the dependent variable as a binary one (voiced vs voiceless) and the second considering the dependent variable as continuous (% of voicing).

4. Methodology

4.1. Participants

A total of 11 speakers from the Basque city of Vitoria were recorded, 5 females and 6 males. One of the male speakers was later rejected for the analysis, since he had spent many of his teen years in multiple Spanish cities. The rest of the speakers were born in Vitoria and had lived in the city at least until they were 18.

The age of the speakers ranged from 24 to 49. 9 speakers reported knowledge of Basque, and one speaker reported being bilingual. 7 speakers reported knowledge of English, although one of the speakers who did not report knowledge of English had worked as an English instructor and had lived

---

3 Hualde suggested that he would not expect any influence of Basque in the speakers from Vitoria, since most of them have Basque as their L2 (personal communication with Hualde on November 23rd 2011 via e-mail)
4 Additional background information on the speakers’ level of Basque and the amount of contact with Basque language would be helpful in determining the potential role of contact. However, this background information was not collected, so these questions will have to be reserved for future research.
in the US for a year. One informant also reported knowledge of German and another reported knowledge of French (her mother was French).

4.2. Data elicitation
4.2.1. Recording

The recordings were collected in June 2011 in Vitoria (9 speakers) and Madrid (1 speaker). The recordings occurred in different places. 3 of them were realized in private residences, 2 in classrooms and 5 in a soundproof booth at the Facultad de Letras of the University of the Basque Country in Álava.

For the recording a Shure WH20XLR head-mounted dynamic microphone was used, attached to a Sony Hi-MD Walkman MZ-RH1. The recordings were then transferred to a Mac computer for an analysis with Praat (Boersma & Weenik, 2010).

4.2.2. Tasks

The speakers had two linguistic tasks. Before the tasks, there was a brief interview so that the participants became more comfortable talking into a microphone and to being recorded. After the interview, the speakers completed two elicitation tasks, created by Schmidt and Willis (2011) and slightly modified to be more appropriate for the dialect being analyzed.

The speakers were first trained on the target vocabulary with the help of orthographic prompts. The speakers were presented with a PowerPoint presentation that showed different pictures, along with some of the letters of the words represented to help the speakers identify the items. For example, the speakers would see the picture of 4 lemons as well as "C U A _ _ _ L I _ _ _ _ _", and they had to say Cuatro limones (‘four lemons’). The goal of this training session was to allow the participants to become familiarized with the lexical items that were going to be elicited during the actual experiment and, thus, to prevent them from using items different than those expected.

The actual experiment was an information gap task where speakers were told that they entered into a fictitious Mercado Santiago and had to say what each vendor in the market sold. They were given the picture of a vendor as well as his/her name and a picture of the item(s) the vendor sold. For example, they saw the picture of Andrés, his name and a picture of 6 cats. The speaker, then, had to say Andrés vende seis gatos (‘Andrés sells six cats’). By this time the speakers had already been trained and they were not shown any orthographic prompt. In this part of the presentation the slides were timed and the speakers only had 5 seconds to produce the utterance, in order to increase the speech rate. To illustrate better the methodology used for the study, three screenshots of the presentation can be found in Appendix B.

As previously mentioned, some items in this task were changed to fit better the dialect under study. For example, words like durazno (‘peach’), tamal, carro (‘car’), and chisme (‘gossip’) were eliminated for the task, and substituted by words like muslo (‘thigh’), tomate (‘tomato’), coche (‘car’), and prismáticos (‘binoculars’) (the full list of contexts is available in Appendix A).

A total of 56 utterances of the sort "(Name of the vendor) vende (‘sells’) (Number of items) (Item)" were produced, although not all the words produced were the ones expected. Some speakers produced words like pastel (‘pie’) instead of postre (‘dessert’), or CD ([/θe’det/]) instead of disco (‘disc’), for example.

The words, quantities and names in this task were included to create phonological contexts leading to s-voicing assimilation. The names of the items sold contained word-internal contexts (prismáticos ‘binoculars’, fantasma ‘ghost’, isla ‘island’, etc.) and the names of the vendors, as well as the number of items sold, helped create contexts of s-voicing assimilation across word boundaries (Nicolás vende dos muslos ‘Nicolas sells two thighs’).

The second task was a reading comprehension task. The participants had to read a text out loud and then answer three questions about the text.

5 The speaker was born and raised in Vitoria and had moved to Madrid a few months before the data collection.
After the two tasks, the participants had to complete a background questionnaire that included questions about the languages they spoke, the native languages of their parents and information about experience abroad.

4.2.3. Coding

In this study, only the tokens in the context [s + voiced consonant] were analyzed, since that is considered to be the context where voicing assimilation is expected. The assimilation process was studied both word internally and across word boundaries. Intervocalic contexts were not analyzed in this study. The tokens were coded for position of the segment (word internally/across word boundaries), place of articulation of the following voiced consonant, manner of articulation of the following voiced consonant, height and frontness/backness of the previous vowel, whether the syllable that contained the /s/ was stressed or unstressed, duration of the /s/ segment, type of voicing for cases of partial voicing (left-edged, right-edged, both), the duration of the voiced part and the percentage of voicing. Finally, the tokens were also coded for the gender of the speaker, since Schmidt & Willis (2011) had found differences between males and females in their data from Mexico, although not statistically significant, the same as Muñiz Cachón and Cuevas Alonso (2012) who also found these differences in Asturian Spanish. At first, those segments with 60% or more voicing were considered voiced, which was the threshold identified by Schmidt and Willis (2011), in order to allow for a cross-dialectal comparison. The authors acknowledged that this cutoff was somewhat arbitrary, but that this way it allowed for replicability and comparison. But then, voicing was taken as a continuous variable for the statistical analysis, since it is a gradual process and analyzing it as continuous can give us more information about the process.

Figure 3. Example of codification for Esmeralda (participant S2).

---

6 One reviewer stated that speech rate could be a factor. However all the speakers participated in the same instrument with the same instructions, therefore a comparable pattern was expected with that respect.

7 The authors identified this threshold by creating a histogram to see the distribution of the tokens and determining whether there was a visible cutoff point. The authors acknowledged that this cutoff was somewhat arbitrary, but that this way it allowed for replicability and comparison (Willis, personal communication).
In order to measure the duration of the /s/ segment, both a change in the formants and a change in the waveform were examined in order to determine where the segment started and ended. The starting and ending points coincided with the zero crossing in order to have some consistency for all tokens. The voiced sections of the segment were determined by the presence of a voice bar, the presence of the periodicity in the waveform and the presence of the F0 with the default settings in Praat. The three correlates were examined before making a decision on where the voiced section started/ended. An example of how the codification was done can be found in Figure 3. There is presence of the voice bar at the beginning of the segment, but the voice bar is interrupted until the following segment /m/. At the same time, in the oscillogram, there is presence of periodicity in the soundwave also at the beginning of the segment, but during the second part of the segment the periodicity disappears. Finally, although not shown in the figure, the first part of the segment had F0, but not the second part. These correlates indicate that the segment was partially voiced. After measuring the segment and measuring the voiced sections, the voicing rate was calculated by multiplying the duration of the voiced part by 100 and then dividing it by the total duration of the segment, thus giving us the voicing rate in the segment.

5. Results

In this section we will present the results of the analysis. The results have been divided into six sections: general results, where we present the global results and the results by gender; the results according to place and manner of articulation of the following segment; the results according to the previous vowel; the results according to stress and position in the word; a description of the position of voicing within the segment (section of the segment were voicing was found), the results from the statistical analysis; a comparison of the results from the current study and the results found in Schmidt & Willis (2011) for Mexico; and, finally, preliminary evidence of cases of devoicing of the following consonant instead of /s/ voicing. In Section 5.1 results are presented based on the two approaches to the variable (binary and continuous), the first to allow for a cross-dialectal comparison with Mexican Spanish (Schmidt & Willis 2011) and, the second because voicing being a gradual process considering degree of voicing as a continuous variable gives us a more clear picture of the process.

5.1. General results

A total of 568 tokens were analyzed. These did not include cases in a [s + rhotic] context, because in this context there was a variety of realizations that included coarticulation resulting in an assimilated rhotic, elision of the /s/ segment and even small pauses between the /s/ and the rhotic. Tokens across word boundaries where there was a brief pause between the /s/ and the following segment were also eliminated from the study.

In the analysis it was found that 44.72% (N=254) of the tokens presented 60% or more voicing, and 55.28% (N=314) presented less than 60% voicing. Table 2 shows the distribution according to gender. Gender was included in the analysis since Schmidt & Willis (2011) had found some differences, although not statistically relevant, between males and females and we wanted to verify if the same was true for the data from Vitoria. As the table shows, male speakers showed more voicing than female speakers. In fact, whereas the majority of the /s/ realizations were voiceless in the case of females, the voiced realizations were the majority in the case of males.

These results will allow us for a comparison with the results from Mexico (Schmidt & Willis 2011) in Section 5.7.

---

8 Although all three correlates were taken into consideration (periodicity of the soundwave, voice bar and F0), not all three were always present and F0 was found to be the least reliable. In the spectrograms included in this paper I have, therefore, not included F0. An additional figure has been added to Appendix C showing both F0 and glottal pulses (as suggested by one of the reviewers).

9 Originally the context [s + rhotic] was going to be coded for since it is a [s + voiced consonant] context. However, given the absence of a clear sibilant fricative without a pause afterwards in the first tokens analyzed, a decision was made not to code that context.
Table 2. Results by gender.

<table>
<thead>
<tr>
<th></th>
<th>N</th>
<th>Males</th>
<th>Females</th>
</tr>
</thead>
<tbody>
<tr>
<td>Voiced</td>
<td>253</td>
<td>52%</td>
<td>146</td>
</tr>
<tr>
<td>Voiceless</td>
<td>315</td>
<td>48%</td>
<td>135</td>
</tr>
<tr>
<td>TOTAL</td>
<td>568</td>
<td>49%</td>
<td>281</td>
</tr>
</tbody>
</table>

Figure 4. Average voicing among males and females

On the other hand, if we take voicing as a continuous variable and we calculate the average degree of voicing of the segments, the average voicing in preconsonantal position was 58%. That is, /s/ before consonants were on average 58% voiced. Of 568 tokens analyzed in the study, less than 50% were fully voiced (41.7%, N=237). The rest was either partially voiced (53.4%, N=303) or completely voiceless (4.9%, N=28). Additionally, we can observe that the average percentage of voicing for males was 62%, whereas the average voicing for females was 54% (Figure 4). Therefore, if we consider voicing as a continuous variable, differences are still found between males and females.

Figure 5. Results according to place of articulation.
5.2. Assimilation according to place and manner of articulation of the following segment

In this section, we present the results depending on the place and the manner of articulation. For place of articulation, there were 6 categories: bilabial, dental, alveolar, palatal, velar and labiovelar. In Figure 5 we can see that, whereas the distribution is rather similar for bilabials, dentals, palatals, velars and labiovelars, ranging from 56% to 65% of average voicing, the average for the alveolars is lower, 42%. It is important to notice that in this category we have the sounds /n/ and /l/, which, as we will see in Table 5, individually also present low rates of assimilation. Each of the other categories represents only one sound. For example, the category ‘Dental’ only includes the sound /d/.

As for manner of articulation, the presence of voicing was examined in three independent variables: approximants (that included the approximants [β], [ð] and [ɣ], as well as the glides [j] and [w]) (N=370), nasals (N=146) and laterals (N=52), which were the types of consonants following the /s/ present in the instrument. The distribution of the results is in Figure 6. It should be noted that in this article we use the term ‘approximants’ when referring to /b d g/, but in a general sense, since no acoustic analyses were done to determine the degree of friction of the consonant (more approximant or more fricative - there were no occlusive realizations in the contexts analyzed).

*10 Those results marked with an asterisk were found as statistically significant in the subsequent analysis (see section 5.6)
As we can see in the distribution, rates of voicing assimilation for nasals and laterals were low, compared to those of the approximants. In light of these results, we selected some of the voiced consonants to see the distribution of the voicing assimilation individually. The selected sounds were [β] (N=162), [ð] (N=17), [ɣ] (N=12), [m] (N=110) and [n] (N=36). [l] was not included, since the results for this sound are already found in Figure 6. The distribution is shown in Figure 7.

The results shown in Figure 6 and Figure 7 show that the assimilation rates for nasals and laterals are lower than for approximants and, with respect to the approximants, that the further back the approximant is produced, the lower the average voicing.

5.3. Assimilation according to vowel height and frontness/backness of the previous vowel

Differences in the voicing rate were found depending on the height and frontness/backness of the previous vowel. The average voicing was higher after mid vowels (61%, N=364) than after high (53%, N=131) and low (46%, N=63) vowels (Figure 8). With respect to the frontness/backness of the vowel, the average voicing was higher when the /s/ was preceded by a back vowel (65%, N=194), then a front vowel (55%, N=301) and then a central vowel (46%, N=63) (Figure 9).

Figures 8 and 9. Results based on vowel height and vowel frontness/backness, respectively.
5.4. Assimilation according to stress and position in the word

As can be observed in Figure 10 small differences were found depending on stress, with a higher voicing average when the /s/ appeared in an unstressed syllable (60%, N= 428; vs. 51%, N=130). Despite the original predictions, this difference was found to be statistically significant (see Section 5.6). As for differences based on the position in the word, more voicing was found when the /s/ appeared word finally (62%, N=412) than when it did word internally (46%, N=146), as it was predicted (Figure 11). However, this difference was not statistically significant.

Figure 10. Results based on stress*.

![Figure 10](image)

Figure 11. Results based on the position of /s/ in the word.

![Figure 11](image)

5.5. Position of voicing occurrence within the segment

Besides looking at the degree of voicing of each segment, we also looked at where the voicing was found in those cases where there was partial voicing. This was done in order to verify whether all were cases of regressive voicing or, on the other hand, whether there were cases of progressive voicing. As can be seen in Figure 12, of all the cases where there was partial voicing (N=302), there were more cases of voicing present at the beginning of the segment (N=178) than at the end of the
segment (N=5), although a high number of tokens also corresponded to cases were there was voicing both at the beginning and at the end of the segment, but not in the middle of the segment (N=119).

*Figure 12.* Position of the voiced section of the segment in tokens with partial voicing.

![Position of voicing](image)

However, if we then look at the degree of voicing depending on where in the segment the voiced part is, it can be observed that there was more voicing on average when it occurred at the end of the segment (29%) than when it occurred at the beginning of the segment (25%) (Figure 13).

*Figure 13.* Degree of voicing of the segment based on where voicing occurred within the segment.

![Degree of voicing](image)

5.6. Results from the statistical analysis

A linear mixed model analysis was performed with the statistical package SPSS in order to determine which factors are significant in the voicing assimilation process. The dependent variable was the voicing percentage and the factors included in the analysis were subject (random effect), position in the word, manner of articulation, place of articulation, stress, and gender of the speaker. The data from the previous vowel was not included given that the type of coding used did not allow for a good pairwise comparison between the factors related to the vowel (vowel /a/ was both the only central and the only low vowel). Additionally, including the five vowels separately was not done either given the low number of cases of /s/ after /w/ (N=22, from a total of 568 tokens). The duration of the
segment was included as a covariate. The data were /s/ was followed by a palatal sound (essentially /j/) were eliminated from the statistical analysis given that there were only 10 cases (out of a total of over 500 tokens). Among all the factors included in the analysis, three were found to be statistically significant: manner of articulation ($p < .05$), stress ($p < .05$) and duration of the segment ($p < 0.1$). As can be seen in Table 3, the results for voicing before approximants were significantly different than those for voicing before nasals and laterals, meaning that the /s/ behaves differently before approximants than before nasals and laterals. As for stress, there were significant differences depending on whether the /s/ was in a stressed or in an unstressed syllable.

**Table 3. Results of the Pairwise Comparisons.**

<table>
<thead>
<tr>
<th>Factor</th>
<th>Comparison</th>
<th>Mean Difference</th>
<th>Sig.</th>
</tr>
</thead>
<tbody>
<tr>
<td>Manner of articulation</td>
<td>Approximants – Lateral</td>
<td>8.704</td>
<td>.046</td>
</tr>
<tr>
<td></td>
<td>Approximants – Nasals</td>
<td>5.704</td>
<td>.018</td>
</tr>
<tr>
<td></td>
<td>Nasals – Lateral</td>
<td>2.941</td>
<td>.430</td>
</tr>
<tr>
<td>Stress</td>
<td>Unstressed – Stressed</td>
<td>3.900</td>
<td>.047</td>
</tr>
</tbody>
</table>

With respect to the duration of the segment and its effect on voicing, a bivariate correlation analysis was performed in order to confirm that the voicing rate was correlated with the duration of the segment. For the analysis, the Spearman Correlation Coefficient was used, given that the distribution of the data was far from perfect. As can be observed in the results of this analysis, presented in Table 4, there was a negative correlation between the voicing rate and the duration of the segment, meaning that the longer the duration of the segment, the lower the voicing rate.

**Table 4. Results of the bivariate correlation between % voicing and duration of the segment.**

<table>
<thead>
<tr>
<th>Correlations</th>
<th>Duration of the segment</th>
<th>% voicing</th>
</tr>
</thead>
<tbody>
<tr>
<td>Spearman’s rho Duration of the segment</td>
<td>1.000</td>
<td>-354**</td>
</tr>
<tr>
<td></td>
<td>Sig. (2-tailed)</td>
<td>.000</td>
</tr>
<tr>
<td></td>
<td>N</td>
<td>568</td>
</tr>
<tr>
<td>% voicing Correlation Coefficient</td>
<td>-354**</td>
<td>1.000</td>
</tr>
<tr>
<td></td>
<td>Sig. (2-tailed)</td>
<td>.000</td>
</tr>
<tr>
<td></td>
<td>N</td>
<td>568</td>
</tr>
</tbody>
</table>

**. Correlation is significant at the 0.01 level (2-tailed).**

5.7. Cross-dialectal comparison

As can be seen in Table 5, the overall assimilation rate was higher for Mexican Spanish than for Basque Spanish by almost 20%, making the assimilation process the majoritary option in Mexico, but not in the Basque Country, at least if we take the DV as a binary one. Moreover, similar to what Schmidt & Willis (2011) found for Mexico, males showed a higher assimilation rate than females, although the differences between males and females were larger in the current study than in the study in Mexico.
Table 5. Comparison of voicing assimilation rates across dialects

<table>
<thead>
<tr>
<th></th>
<th>Schmidt &amp; Willis (2011)</th>
<th>Current study</th>
</tr>
</thead>
<tbody>
<tr>
<td>Overall assimilation rate</td>
<td>63%</td>
<td>44.7%</td>
</tr>
<tr>
<td>Distribution according to sex</td>
<td></td>
<td></td>
</tr>
<tr>
<td>Males</td>
<td>67%</td>
<td>52%</td>
</tr>
<tr>
<td>Females</td>
<td>59%</td>
<td>37%</td>
</tr>
<tr>
<td>Difference</td>
<td>8%</td>
<td>15%</td>
</tr>
</tbody>
</table>

5.8. Devoicing

Finally, in the data from the current study there were also some instances in which, instead of having a voicing assimilation process, we observed devoicing of the following segment (Figure 14). In the case of Figure 4, the sound /g/ was produced as a voiced velar fricative ([x]).

*Figure 14. Devoicing of the segment following the /s/ in unas gafas ‘some glasses’ (found in 2 speakers)*

This devoicing process was observed in at least four different speakers and, in all cases, before an approximant. As it was previously mentioned (see section 2.2), there is a phonological process in
Basque that consists of the devoicing of voiced stops following a voiceless consonant (Hualde, 1991). However, it is important to bear in mind that the number of cases in this corpus is very limited and that cases of devoicing were also found in Asturias (Muñiz Cachón & Cuevas Alonso 2012). So, at this point we cannot theorize on whether Basque has any influence on this process.

6. Discussion

The results from the present study indicate that 45% of the tokens in preconsonantal position were 60% or more voiced and that, in this position, the average voicing rate was 58%. Additionally, less than half the tokens were completely voiced, the most frequent realization being that of partial voicing. It was also observed that the assimilation rate varied between males and females. This difference was not statistically significant, probably due to the fact that the differences between males and females laid in a different distribution of the data. That is, whereas the distribution of the rate of voicing among females was somehow more balanced, males tended to have a more polarized distribution, with more cases of either fully voiced or fully voiceless tokens than females. Muñiz Cachón and Cuevas Alonso (2012) offer two possible explanations to these differences between male and female speakers. First, they mention physiological factors, also mentioned by Nadeu and Hualde (2013). Nadeu and Hualde (citing Lucero & Koenig, 2005) explain that a larger larynx for the males might result in a higher difficulty to control voicing/voicelessness at least in intervocalic position. However, Muñiz Cachón and Cuevas Alonso (2012) disregard this hypothesis stating that both males and females are capable of producing voiced and voiceless /s/ and the distribution of the tokens in the present study seems to support this position. Another explanation offered by Muñiz Cachón and Cuevas Alonso (2012) for these differences is a potential role of sociolinguistic factors. The authors argue that females have a more conservative tendency and that they show “realizaciones más tensas, menos perezosas, poco propensas al contagio de la sonoridad” (298) (more tense, less lazy realizations, less prone to voicing transmission” (my translation). It is unclear, though, what they mean by lazy realizations. Finally, Nadeu and Hualde (2013), talking about voicing of intervocalic /p t k/, where differences were also found between males and females, mention the possibility that speech rate might play a role, too. The negative correlation found between the duration of the segment and the amount of voicing present in the segment (which was statistically significant) seems to confirm this. Faster speech results in shorter segment durations and, at the same time, shorter segment durations show on average higher voicing rates. This coincides with García's (2013) findings, who found differences according to speech rate in the production of intervocalic /s/ in Highland Colombian Spanish speakers and of preconsonantal /s/ in a group of speakers from different Spanish-speaking areas. However, in Nadeu and Hualde’s (2013) study speech rate does not seem to be the only factor playing a role in intervocalic stop voicing. It seems, then, that several factors combined, rather than separately, would explain the differences according to gender; physiological differences, speech rate and sociolinguistic factors, although these last two might be closely related in the sense that speech rate might be subject to sociolinguistic factors. In any case, it seems clear that speech rate is a factor that needs to be incorporated or, at least, controlled for in future studies.

In regards to the research questions and hypotheses that guided this study, we can observe that not all hypotheses were confirmed. To the question of whether the voicing rate was the same before all consonants, the data show that it varied depending on both the manner and the place of articulation, as it was expected given the results for Asturian Spanish (Muñiz Cachón and Cuevas Alonso 2012). If we compare the rates depending on place of articulation, we can observe that there was less voicing when the following consonant was alveolar than before any other place of articulation. On the other hand, there was more voicing before approximants than before nasals and laterals. The effect of manner of articulation was statistically significant. This effect of the manner of articulation might be what explains why there was less voicing before alveolar sounds than before any other sound, given that in the data essentially all alveolar consonants were either /n/ or /l/. A potential explanation for the significance of manner of articulation would be differences in the magnitude of the articulatory gesture. According to Browman & Goldstein (1989) those consonants and sequences of consonants that require a larger magnitude of gesture will be less prone to coarticulation. It is possible that the sequence /s/ + approximant might show a smaller articulatory gesture than the sequence /s/ + lateral or
nasal, which might lead to a higher coarticulation of the glottal configuration in those cases where the sequence is /s/ + approximant.

To the question of whether the assimilation rate varies depending on the previous vowel we can see that the /s/ segment shows on average more voicing more frequently after mid and back vowels and less voicing after back and central vowels (/a/). However, given that the consonant following the /s/ was not kept constant and that the distribution of tokens depending on the previous vowel was not balanced, it is difficult to determine to what extent the previous vowel plays a role. In future studies the phonetic context (both previous and following) should be controlled to allow for a detailed analysis of the role of the different contexts.

To the question of whether there were cross-di alectal differences, we can observe that the assimilation rates were different in Mexican Spanish and in Basque Spanish. It is still unclear why there are such differences between the dialects, but there are two possible explanations for these differences. First, although there is limited contact between Spanish and Basque in Vitoria and it has been claimed that the assimilation process exists in Basque, we cannot completely rule out a possible influence of Basque. Some reasons for not ruling out this possibility yet are that there is little research on the assimilation process in Basque, so we do not know how it compares to the process in Spanish, and that there is an opposite process of devoicing in Basque. Some instances of a devoicing process were found in this corpus, which might or might not be due to the influence of Basque, but it is worth analyzing more in detail in future studies. Further research is needed to verify if there really is an influence of Basque, including a comparison between the data from this study and data from other northern monolingual Peninsular varieties, as well as a comparison between speakers with different levels of exposure to Basque. If differences are found, we could then explore the hypothesis of the influence of Basque. If, on the other hand, similar results are found, there might be a second possible explanation for the results, other than language contact. A potential difference in the quality of the /s/ might be behind the differences between the results from Mexico (Schmidt & Willis 2011) and Vitoria, with some realizations of /s/ favoring the voicing assimilation more than others for articulatory reasons. Again, more comparative studies that include acoustic analysis of the quality of /s/ and assimilation rates in different dialects would be needed to understand better the role of Basque and the role of the quality of the /s/ in the voicing assimilation process.

With respect to the fourth question, differences were found depending on the position in the word (more voicing on average word finally than word internally) and on the stress of the syllable where the /s/ is found (more voicing in unstressed syllables). This last factor was statistically significant. The differences based on the stress of the syllable might be connected with the length of the segment. Unstressed syllables tend to be shorter than stressed syllables (and, arguably, the different segments within the syllable could be subject to the same effect) and the results from this study show a negative correlation between the duration of the segment and the degree of voicing. So it is possible that the /s/ in unstressed syllables are shorter than in stressed syllables and this would explain why more voicing is found in unstressed syllables.

Finally, the distribution of the presence of voicing within the segment shows presence of voicing not only at the end of the segment, prior to the voiced consonant, but also at the beginning of the segment, which suggests coarticulation with the previous vowel, something already suggested by Schmidt and Willis (2011) for their data from Mexico.

7. Conclusion

As the data in this study have shown, /s/ voicing assimilation is not as systematic as it may appear to be. It is not equally favored in all the possible contexts and it is subject to variation depending on different factors such as stress, position in the word, duration of the segment and gender of the speaker. Additionally, it is not solely a regressive process. As the data has shown, there is also presence of coarticulation with the previous vowel. Additionally, dialectal differences were also found as, after a comparison with the results by Schmidt & Willis (2011) in Mexico, it could be observed that overall the assimilation rate was higher for Mexican Spanish, although it remains to be analyzed if these differences are due to differences in the quality of /s/ in both dialects, if the lower assimilation rate in
the Spanish of the Basque Country is due to contact with Basque or if there is another reason that would explain these differences.

8. Future directions

The results from the current study have shown that the voicing assimilation rule is not completely systematic, but, rather, a variable rule that depends on both linguistic and extralinguistic factors. However, further research is still needed to better understand how the assimilation process actually works how the assimilation process works in contact situations. In this sense, several future directions are possible. First, comparing the current data with data from monolingual regions around the Basque Country and in northern Spain in general in order to verify a potential influence of Basque. Controlling for knowledge of Basque or exposure to the language would also help in this sense. Additionally, a sociolinguistic study could be carried out in order to analyze if social factors other than gender have an effect on the application of the process. More detailed research is needed, as well, to determine the role of the different factors in the voicing rate, but controlling for a balanced distribution of the tokens depending on the different factors and, that way, confirm the findings in the present study.

Appendix A
List of tokens elicited

<table>
<thead>
<tr>
<th>Across word boundaries</th>
<th>Word-internally</th>
</tr>
</thead>
<tbody>
<tr>
<td>Dos bolsos</td>
<td>Dos mesas</td>
</tr>
<tr>
<td>Carlos vende (x2)</td>
<td>Seis nueces</td>
</tr>
<tr>
<td>Dos rosas</td>
<td>Dos hierbas</td>
</tr>
<tr>
<td>Mercedes vende (x2)</td>
<td>Tres huevos</td>
</tr>
<tr>
<td>Tres vestidos</td>
<td>Jugadores de</td>
</tr>
<tr>
<td>Andrés vende (x2)</td>
<td>Nicolás vende (x2)</td>
</tr>
<tr>
<td>Luis vende</td>
<td>Tomás vende (x2)</td>
</tr>
<tr>
<td>Dos máscaras</td>
<td>Dos bisnietos</td>
</tr>
<tr>
<td>Tres limones</td>
<td>Dolores vende (x2)</td>
</tr>
<tr>
<td>Seis vasos</td>
<td>Marcos vende</td>
</tr>
<tr>
<td>Cubitos de</td>
<td>Milagros vende (x2)</td>
</tr>
<tr>
<td>Seis gatos</td>
<td>Dog muslos</td>
</tr>
<tr>
<td>Inés vende (x2)</td>
<td>Seis dientes</td>
</tr>
<tr>
<td>Unas gafas</td>
<td>Seis narices</td>
</tr>
<tr>
<td>Dos botas</td>
<td>Dos ranas</td>
</tr>
<tr>
<td>Dos globos</td>
<td>Dos ratones</td>
</tr>
<tr>
<td>Dos lápices</td>
<td>Tres huesos</td>
</tr>
<tr>
<td>Tres libros</td>
<td>Tres huellas</td>
</tr>
<tr>
<td>Tres discos</td>
<td>Israel (x2)</td>
</tr>
<tr>
<td></td>
<td>Esmeralda (x2)</td>
</tr>
<tr>
<td></td>
<td>Rosbeef</td>
</tr>
<tr>
<td></td>
<td>Esmalte</td>
</tr>
<tr>
<td></td>
<td>Fantasma</td>
</tr>
<tr>
<td></td>
<td>Ismael (x2)</td>
</tr>
<tr>
<td></td>
<td>Béisbol</td>
</tr>
<tr>
<td></td>
<td>Isla</td>
</tr>
<tr>
<td></td>
<td>Prismáticos</td>
</tr>
<tr>
<td></td>
<td>Muslos</td>
</tr>
<tr>
<td></td>
<td>Asnos</td>
</tr>
<tr>
<td></td>
<td>Cisne</td>
</tr>
</tbody>
</table>
Appendix B

Screenshots of the slides in the PowerPoint Presentation. The first two correspond to the training session, whereas the last two correspond to the actual experiment.

Sesión de práctica
Antes de completar una actividad en el Mercado Santiago, realizarás una prueba rápida para familiarizarte con las mercancías que encontrarás en el mercado.

Instrucciones: Di en voz alta lo que ves en cada dibujo que aparece a continuación. Después de describir lo que ves, haz clic para seguir al siguiente dibujo.

¡Bienvenido al Mercado Santiago!

Instrucciones: Es el sábado por la mañana, y caminas por el Mercado Santiago. Estás encargado de reportarle por teléfono a la señora Gómez (su vecina entrometida) lo que vende cada vendedor(a) en el mercado.

¡Cuidado! La señora Gómez tiene poca paciencia. Por eso tienes que hablarle rápidamente o ella pierde interés. ¡(Tienes sólo 5 segundos para describir lo que vende cada vendedor!)}

¿Qué vende?

don Israel
Appendix C

Example of codification for Esmeralda (participant S2).

References


Torreblanca, Máximo. 1978. La ‘S’ sonora prevocálica en el español moderno. Thesaurus, XL1 (1, 2 y 3), 59-69.