Phonological Retention and Innovation in the Judeo-Spanish of Istanbul

Travis G. Bradley and Ann Marie Delforge
University of California, Davis

1. Introduction

The term “Judeo-Spanish” denotes the varieties of Spanish spoken by the Sephardim, descendents of Jews expelled from Spain in 1492. While the Sephardim who settled in Italy, the Low Countries, and other parts of Europe ceased to speak Spanish soon after the Expulsion, those who immigrated to Turkey, Greece, the Balkans, and Northern Africa preserved the language they carried with them from Spain for over four hundred years. This remarkable conservation of a non-national language was facilitated by the policies of the Ottoman Empire which afforded considerable autonomy to its minority populations (Diaz Mas 1997, Harris 1994). Sephardic groups living in Ottoman territories functioned as virtually self-governing communities with their own neighborhoods, schools, and synagogues and were thus able to retain Judeo-Spanish (henceforth, JS) as their primary language. The Sephardim who settled in Morocco were also allowed the independence necessary to preserve their language, although they did not enjoy the same privileges and level of prestige as their counterparts residing in the Ottoman Empire (Diaz Mas 1997).

Unfortunately, the circumstances that had favored the conservation of JS began to change during the nineteenth century. In Morocco, the language of the Sephardim lost most of its unique characteristics due to contact with modern Spanish brought about by the military and political presence of Spain in the area starting in the 1860s. The rise of nationalism in the Balkans during the late 1800s and the subsequent dissolution of the Ottoman Empire in the early twentieth century undermined the conditions that had permitted the maintenance of JS in Ottoman territories for so many years. New policies imposing the use of national languages and requirements that all children attend state-run schools impeded the transmission of JS to younger generations. Also, popular schools founded by the Alliance Israélite Universelle, a group dedicated to the westernization of Sephardic Jews, induced many young Sephardim to adopt French as their first language. Large numbers of Sephardim emigrated to Israel and the U.S. in order to escape the political and economic instability of the region, and their children learned Hebrew or English rather than JS. And, sadly, the Second World War resulted in the death of many who remained in the Balkans. As a consequence of these various adverse circumstances, JS is currently considered a dying language (Diaz Mas 1997, Harris 1994, Malinowski 1982). Harris (1994) estimates that there were only sixty thousand proficient speakers of JS in the early 1990s. At that time, there were no longer any monolingual speakers and very few fluent speakers under the age of fifty-five.

Despite its endangered status, JS has been underrepresented in the most recent work on Ibero-Romance phonetics and phonology. The present study, an examination of two salient aspects of JS pronunciation based on data from an elderly female speaker from Istanbul, represents an initial attempt to remedy this neglect. Some previous studies (e.g., Lapesa 1980:526) have emphasized the conservative, archaic nature of JS, highlighting its similarities to fifteenth century Castilian Spanish. Other descriptions (Diaz Mas 1997, Harris 1994, Penny 2000) emphasize the novel characteristics that

* For comments on an earlier presentation of this work, we acknowledge the audience members of the Eighth Hispanic Linguistics Symposium, held in October, 2004, at the University of Minnesota. We also thank Barbara Bullock, John Lipski, Orhan Orgun, and Kimball Robinson, as well as two anonymous reviewers. We alone are responsible for any errors or shortcomings in this paper.
distinguish JS from both medieval Castilian and other contemporary Spanish varieties. In keeping with the dual focus of previous research, we analyze two conservative characteristics of JS, namely the retention of the medieval sibilant voicing contrast and the voicing of word-final intervocalic sibilants, as well as one of its innovative features, namely the loss of the contrast between the rhotic tap /ɾ/ and trill /r/. First, our results show that the distinction between voiced and voiceless sibilants is maintained word-internally but that voicing in word-final intervocalic position is more variable than has been indicated in previous descriptions of JS. The difference observed between the realizations of sibilant voicing in word-medial and external sandhi contexts is interpreted as supporting a phonetic underspecification approach to obstruent voicing neutralization (Ernestus 2003, Steriade 1997). In this approach, obstruents that are phonologically contrastive in voicing are distinguished from neutral obstruents, which are targetless with respect to [voice] in the output of phonology. Neutral obstruents are subject to variable and gradient voicing in the phonetic implementation component, due to the interpolation of glottal activity from the surrounding context. Second, we present a qualitative, acoustic description of syllable-initial rhotics, showing that the multiple trill /r/ is absent from the speech of our Istanbul JS informant. Rather, a syllable-initial rhotic in any position is a voiced approximant [ɹ] of variable duration, which often tends toward frication. Frication and devoicing of word-final rhotics is also observed in prepausal contexts, most likely due to phonological transfer from Turkish. Our descriptive results are in agreement with Hammond’s (1999, 2000) findings for many dialects of contemporary Spanish, namely that the prescriptive multiple trill simply does not occur in normal discourse of the vast majority of speakers.

The remainder of this paper is organized as follows. Section 2 provides some historical background on the sibilants, tracing their development from medieval Castilian through contemporary varieties including JS. Section 3 details the data collection and analysis procedures employed in our study of Istanbul JS sibilants. Section 4 documents the observed sibilant voicing patterns and situates them within a theoretical discussion of obstruent voicing neutralization. Section 5 presents an acoustic description of rhotics and relates our findings to rhotic patterns in contemporary Spanish varieties. Section 6 summarizes and concludes.

2. Sibilant voicing contrasts in medieval Castilian and JS

2.1 The medieval Castilian sibilants

The phonemic inventory of modern Spanish contains only two sibilants, both of which are voiceless: /s/, which is apicoalveolar in Northern Spain and dorsoalveolar in most other areas, and /ʃ/, a voiceless prepalatal affricate. However, the medieval Castilian consonant system possessed a much larger group of sibilants including not only /ʃ/ but also three cognate pairs: two dental affricates (1a), two apicoalveolar fricatives (1b), and two prepalatal fricatives (1c). These voiced and voiceless sibilants contrasted only in intervocalic position as shown in the following examples, taken from Penny (1993:82).

<table>
<thead>
<tr>
<th>(1)</th>
<th>Voiceless</th>
<th>Voiced</th>
</tr>
</thead>
<tbody>
<tr>
<td>a.</td>
<td>detˈir</td>
<td>dedˈir</td>
</tr>
<tr>
<td></td>
<td>deˈcir</td>
<td>dezir</td>
</tr>
<tr>
<td>b.</td>
<td>foˈtes</td>
<td>foˈdes</td>
</tr>
<tr>
<td></td>
<td>foˈxes</td>
<td>fozes</td>
</tr>
<tr>
<td>c.</td>
<td>koˈfo</td>
<td>koˈjo</td>
</tr>
<tr>
<td></td>
<td>koˈxo</td>
<td>kojo</td>
</tr>
<tr>
<td>d.</td>
<td>fiˈfo</td>
<td>fiˈjo</td>
</tr>
<tr>
<td></td>
<td>fiˈxo</td>
<td>fijo</td>
</tr>
</tbody>
</table>

The voicing status of sibilants in coda position was not reflected in the orthography of medieval Castilian, with the graphemes <>, <>, and <> being employed to represent all six phonemes. While we cannot be certain how these sounds were actually realized, Penny (1993) has proposed several hypotheses based on data from modern Spanish as well as from Catalan, Portuguese and JS. He suggests that syllable-final preconsonantal sibilants exhibited regressive voicing assimilation, taking on the voicing value of the following consonant within the same word or across word boundaries, just as /s/ has been observed to do in contemporary Spanish, shown in (2a,b). Presumably, such voicing
was gradient and varied depending on stylistic factors as is the case in the modern language (Navarro Tomás 1977).

(2) a. des\øde \sim dezøe \quad desde \quad 'since'
b. las\βakas \sim laz βakas \quad las vacas \quad 'the cows'

Penny (1993) also argues that, although the /s/ of contemporary standard Spanish remains voiceless in word-final intervocalic position, the medieval Castilian sibilants would have become voiced in this context as do sibilants of Portuguese, Catalan and JS. Examples of this hypothetical voicing of word-final intervocalic sibilants are provided in (3).

(3) laz alas \quad las alas \quad 'the wings'
az iðo \quad has iðo \quad 'you've gone'

A variety of changes affected the sibilant pairs shown in (1) during the medieval and early modern periods. The earliest of these processes was most likely the deaffrication of the dentals /t\üstü/ and /d\üstü/ which probably began in the mid-thirteenth century (Harris-Northall 1992) and was completed by the fifteen hundreds (Lantolf 1974), producing the dental fricatives /s/ and /z/. The merger of the voiced sibilants /z/, /z/, and /ʒ/ with their voiceless counterparts may also have begun at a very early date, perhaps as soon as the twelfth century, in the northernmost areas of Old Castile but spread southward in a significant fashion only during the fifteenth and sixteenth centuries as a result of the political power and prestige acquired by this region during the Reconquest. The voiced sibilants were finally eliminated in even the most linguistically conservative Castilian-speaking areas by the 1580s (Lantolf 1974, 1979). The results of deaffrication and devoicing are represented in (4a,b).

(4) a. det\ir \rightarrow de\ðir \quad dezir \quad deçir \quad 'to descend'
ded\ir \rightarrow dezir \quad deçir \quad 'to say'
b. dezir \rightarrow dezir \quad dezir \quad 'to say'
espezo \rightarrow espeso \quad espeso \quad 'spent'
kəço \rightarrow kojo \quad cojo \quad 'I grasp'

Additional developments took place during the sixteenth and seventeenth centuries. In Northern Spain, the dental fricative /s/ developed an interdental, non-sibilant articulation. In much of Andalusia, the medieval sibilants represented in (1a,b) have merged to a single voiceless dental fricative /s̪/, which may be pronounced with or without interdental coloring, [s̪θ] or [s̪], depending on geographic and social factors (Penny 2000:118-120). In Northern Spain, the prepalatal sibilant /ʃ/ was retracted in place of articulation to a velar fricative /x/. Retraction was more extreme in Andalusia, resulting in what is now a laryngeal or glottal fricative /h/ (Penny 2002:103). These changes are exemplified in (5a,b), along with forms in modern orthography.

(5) a. de\ðir \rightarrow de\ðiɾ \quad deçir \quad 'to say' \quad (Castile)
ozo \rightarrow oso \quad oso \quad 'bear'
kəço \rightarrow koko \quad cojo \quad 'I grasp'
b. de\ðir \rightarrow de\ðiɾ \sim de\ðiɾ \quad deçir \quad 'to say' \quad (Andalusia)
ozo \rightarrow o\ðiɾ ozo \quad oso \quad 'bear'
kəço \rightarrow kojo \quad cojo \quad 'I grasp'

1 The devoicing of the medieval Castilian sibilants has been the subject of numerous publications. For further information regarding chronological issues and potential explanations of the process, see Alonso (1967, 1969), Ford (1900), Martinet (1952), Penny (1993), Pensado (1993) and Widdison (1995, 1996), inter alia.
2.2 Retention of the medieval sibilant voicing contrast in JS

As JS is based on the Castilian spoken during the last decade of the fifteenth century, most of the changes discussed above occurred too late to affect the development of its sibilant system. The very early deaffrication of the dentals /tʰ/ and /dʰ/ is reflected in JS consonant inventory, which includes the fricatives /s/ and /z/. However, the place shifts that produced the interdental /θ/ and the velar /x/ took place over one hundred years after the Expulsion and, while the devoicing process was underway in many areas by the 1490s, the merger of /z/, /z/, and /ʒ/ with /s/, /s/, and /ʃ/ was by no means complete at that time. JS, therefore, essentially retains the sibilant inventory of medieval Castilian with one exception: the apicoalveolars /s/ and /z/ merged with the dentals /s/ and /z/. Thus, the JS sibilant system is identical to that of Central and Southern Portugal. The examples in (6), adapted from Penny (2000:182), contrast the late-medieval Castilian sibilants with the sibilants of contemporary JS.

(6) a. Voiceless Voiceless (Late-medieval Castilian)
   /s/ álcar 'raise' /z/ dezir 'say'
   /s/ passo 'step' /z/ casa 'house'
   /ʃ/ caixa 'box' /ʒ/ muger 'woman'
b. Voiceless Voiceless (Contemporary JS)
   /s/ alsar 'raise' /z/ dezir 'say'
   /ʃ/ casha 'box' /ʒ/ mujer 'woman'

As mentioned previously, JS also exhibits the phonetic voicing of word-final intervocalic sibilants believed to have occurred in medieval Castilian. Penny (2000:182) gives the examples shown in (7).

(7) ma[z] o menos 'more or less'
    do[z] o tres 'two or three'

Based on his description of the phenomenon, however, it remains unclear to what extent this external sandhi voicing actually occurs across different phrasal contexts in contemporary JS.

3. Phonological retention in the sibilant system of Istanbul JS

Given the hypothesis that sibilant voicing is universally disfavored on aerodynamic grounds (Kirchner 1998, Ohala 1983, Widdison 1997), one might expect to find some devoicing in present day Ibero-Romance varieties that have retained contrastively voiced sibilants. This appears to be the case for at least some speakers of contemporary European Portuguese (Jesús & Shadle 2002). In the present study, we carried out an acoustic investigation of the variety of JS spoken in Istanbul in order to determine the status of sibilant voicing in two intervocalic contexts: word-medial (6b) and word-final (7). We constructed the phonetic corpus used in this study with GO, a sixty-seven year old female speaker of JS from Istanbul. Our informant is multilingual, speaking French, Turkish, English and some Greek in addition to JS. While her extensive linguistic repertoire raises the possibility of language transfer, this complication appears to be unavoidable in the study of JS. As was mentioned previously, the language currently has no monolingual speakers (Harris 1994). In any case, it seems that for the Sephardim of Istanbul, multilingualism is a long-standing consequence of the linguistic plurality of the Balkans. Wagner (1930:45, cited in Harris 1994:209) described the linguistic situation of this community in the early twentieth century, observing that “...any youth from a good Levantine family knows how to express himself in at least six languages” with Turkish, Greek, French, Italian, English and German being the languages de rigueur. Thus, while the multilingualism of today’s JS

---

2 In a discussion of the effect of dialect mixing on the characteristics of JS, Penny (1992) suggests that Portuguese-speaking Sephardim may have influenced the development of the language’s sibilant system. Furthermore, see Penny (2000:180, 186) on innovations involving voiced dental and prepalatal affricates in contemporary JS, which we do not examine in the present study.
speakers no doubt reflects the language’s decline and therefore differs from that discussed by Wagner, it can be argued that a number of languages spoken in the Balkans have long influenced the speech of the Sephardim in Istanbul. In addition, the fact that our informant’s linguistic background is very similar to that of the participants in Malinowski’s (1982) study of Istanbul JS supports GO’s assertion that she is a typical speaker of her variety of the language.

One problem facing any phonetic study that involves bilingual or multilingual speakers is the high degree of variability observed in speech production tasks. Controlling for the myriad factors that yield speaker variability in production is notoriously difficult, if not impossible (see, for example, Beach et al. 2001, Bullock et al. 2004, Martinez-Dauden & Llisterri 1990). In a discussion of variation and phonetic convergence in Spanish/English bilinguals, Bullock et al. (2004) observe the following:

This variation may be due to multiple factors beyond those that are inherent in any phonetic production experiment. Factors which include language dominance, speakers’ attitudes about their languages, and language activation, among others, may be so individualistic that it may not, in fact, be possible to control for them all in bilingual production studies ... Given the variation across speakers ..., future studies should focus on the comparison of an individual speaker’s productions across different conditions obviating the need to make generalizations across subjects [our emphasis]. (111-12)

Given the precarious status of contemporary JS and the highly multilingual context in which it exists in Istanbul, a study comparing a single speaker’s productions across different conditions is justified.

Sibilant tokens were obtained by recording GO’s speech in two production tasks. The first elicited the production of single words with intervocalic sibilants, as well as phrases with specific syntactic structures containing word-final <s> before a vowel-initial word (i.e., determiner plus noun, noun adjective combinations, etc.), embedded in the carrier phrase “Diga _____ por favor.” These items were constructed on the basis of the Diksionario de Ladinokomunita, an online Ladino-Spanish-English-Turkish dictionary. In the second task, GO read an article written in JS (approx. 2000 words), taken from an online newspaper, Şalom, published in Istanbul. Both the carrier phrase items and the longer text follow the Latin orthographic system, as most commonly employed in Turkey (see the grapheme-phoneme correspondences in (6b)). Recordings were made in a quiet room on a Sony MZ-R700 mini-disc recorder with a Sony ECM-717 unidirectional microphone. The recordings were subsequently digitized to computer and stored in .wav format at 22,050 Hz and 16-bit.

Tokens of both word-medial intervocalic and word-final intervocalic sibilants were extracted from the digitized corpus and analyzed via waveform and spectrogram using SIL Speech Analyzer software, version 1.5. The length of the sibilant noise interval and the duration of glottal tone were measured manually, and the ratio of glottal tone to sibilant noise was calculated. The resulting percentages were used to divide the tokens into three phonetic categories: voiceless, partially (de)voiced and fully voiced. This classification scheme follows the procedure used in recent studies of sibilant devoicing in American English (Smith 1997) and in European Portuguese (Jesús & Shadle 2002). Inclusion of a “partially (de)voiced” category is based on previous analyses of voiced sibilants in British and American English (Haggard 1978, Stevens et al. 1992). This research indicates that the vocal folds often do not vibrate throughout the production of phonologically voiced sibilants, presumably as a result of the fundamental incompatibility between the aerodynamic demands for the production of sibilant noise and glottal tone (see Section 4 for further details).

Our classification criteria most closely correspond to those used by Jesús and Shadle (2002). Tokens classified as voiceless exhibited glottal tone during less than one third of their frication interval, while partially (de)voiced sibilants displayed periodic structure during more than one third but less than one half of their duration. Tokens having glottal tone during one half or more of their

3 The online dictionary is available at [http://lingua.cc.sophia.ac.jp/diksionario-LK/] and continues to be developed under the direction of Professor Antonio Ruiz Tinoco of Sophia University in Tokyo, Japan. Our informant participates in its development by collecting and contributing lists of JS words and phrases with multilingual translations and examples.

4 The online newspaper is available at [http://www.salom.com.tr].
frication interval were classified as fully voiced. Waveforms and spectrograms exemplifying each of these categories are shown in Figures 1 through 3.

Figure 1: Voiceless sibilant in /empeso/ empeso 'began'

Figure 2: Partially (de)voiced sibilant in /kezo/ kezo 'cheese'
4. Results and discussion of sibilant voicing patterns

Our results confirm that the contrast between voiced and voiceless word-medial intervocalic sibilants is generally maintained in the phonological system of our Istanbul JS informant. The data also demonstrate that word-final intervocalic sibilants are usually voiced, although the significant number of voiceless tokens observed in this context suggests that the effect should be considered a variable rather than categorical phenomenon.

4.1 Word-medial sibilants

Tables 1 and 2 present the frequencies of phonetic realizations of word-medial intervocalic sibilants as produced by our informant in the carrier phrase task. We find that the realization of word-medial sibilants is dependent on the sibilant’s phonological voicing specification, with Pearson’s Chi-square tests indicating significance at p<0.0001 for both dentals and prepalats. In both cases, the most frequently occurring phonetic category corresponds to the sibilants’ phonological specification. We conclude that the distinction between voiced and voiceless sibilants is generally preserved in our informant’s speech. However, for each sibilant, there are a number of tokens that do not conform to the sound’s underlying specification for [voice]. These discrepancies can likely be attributed to several phonetic factors.

<table>
<thead>
<tr>
<th></th>
<th>[-voice] dental</th>
<th>[+voice] dental</th>
</tr>
</thead>
<tbody>
<tr>
<td>Voiceless</td>
<td>49 (68.1%)</td>
<td>2 (2.8%)</td>
</tr>
<tr>
<td>Partially (de)voiced</td>
<td>6 (8.3%)</td>
<td>6 (8.3%)</td>
</tr>
<tr>
<td>Fully voiced</td>
<td>17 (23.6%)</td>
<td>64 (88.9%)</td>
</tr>
<tr>
<td>TOTALS</td>
<td>72 (100%)</td>
<td>72 (100%)</td>
</tr>
</tbody>
</table>

Table 1: Voicing contrast in word-medial intervocalic dental sibilants (words embedded in carrier phrases): $\chi^2$ (df 2, n 144) = 70.59, p < 0.0001
The relatively small number of voiceless and partially voiced tokens of /z/ and /ʒ/ observed in the data are probably a result of a universal, aerodynamically-based tendency for voiced sibilants to exhibit devoicing. As noted above in Section 3, it is generally acknowledged that the production of voiced obstruents, and particularly voiced sibilants, involves conflicting airflow demands. Ohala (1983) observes that intraoral air pressure must be relatively high in order to generate the turbulent noise characteristic of sibilance, while the vocal folds can vibrate only when air pressure in the oral cavity is significantly lower than subglottal pressure. In order to produce voiced sibilants, speakers must make fine motor adjustments ensuring that the cross-sectional areas of the supraglottal and glottal constrictions are approximately the same size. If vocal cord vibration is to be sustained during long intervals of frication, expansion of the pharynx and lowering of the larynx may also be required. Presumably, these articulatory challenges are reflected in the cross-linguistic scarcity of voiced sibilants compared to their voiceless counterparts documented by Maddieson (1984). The observation that phonologically voiced sibilants often exhibit glottal tone during only a portion of their fricative noise is also ascribed to these production difficulties (Haggard 1978, Jesús & Shadle 2002, Smith 1997, Stevens et al. 1992). Thus, it seems likely that the voiceless and partially voiced tokens of /z/ and /ʒ/ are simply a result of the natural predisposition to devoice that characterizes all voiced sibilants.

Given the well-established conflict between voicing and sibilance, it may appear surprising that a small percentage (12.3%) of the phonologically voiceless prepalatal and a more considerable number (31.9%) of the voiceless dental fall into the voiced and partially voiced categories. The explanation of this apparently anomalous finding probably lies in the nature of the phonetic context. Westbury and Keating (1986) as well as Kirchner (1998) note that some degree of articulatory effort is required to produce voiceless obstruents in intervocalic position. When flanked on both sides by voiced sounds, obstruents tend to passively voice in the absence of a specific gesture that abducts the vocal folds. We observed that the voiced portion of phonologically voiceless /s/ and /ʃ/ was most often located at the V-C margin, suggesting that perhaps there was a slight delay in glottal abduction that allowed glottal tone to “seep” into the sibilant at its left margin.

It is interesting to note that many more tokens of /s/ than /ʃ/ were classified as either partially or fully voiced. This difference is reminiscent of the greater propensity of stops with anterior places of articulation to support voicing discussed by Ohala (1983). Ohala argues that phonologically voiced labial and alveolar stops are more common cross-linguistically than voiced velar stops because they are associated with lower intraoral air pressure. The slightly larger oral cavity space created by more forward points of constriction allows supraglottal air pressure to increase more slowly and thus permits vocal fold vibration for a longer period of time. Also, in the case of an anterior constriction, the air in the oral cavity comes into contact with more compliant tissues (the tongue surface, cheeks, and the soft palate) which are more likely to expand, thus further decreasing intraoral pressure. It is reasonable to speculate that the same factors might allow dental /s/ to undergo passive voicing more frequently than prepalatal /ʃ/ in intervocalic position.
4.2 Voiced sibilants in word-medial position versus external sandhi

Table 3 presents results from the text reading task, comparing the realizations of word-medial intervocalic /z/ with those of dental sibilants in external sandhi (i.e., word-final intervocalic position). The Chi-square test indicates statistical significance at p<0.0001, suggesting that the differences observed in phonetic voicing categories depend on whether the sibilant is word-medial or word-final. As in the single word production task shown in Table 1, most tokens (87.9%) of word-medial phonologically voiced sibilants from the reading task fall into the fully voiced category. On the other hand, only 60.3% of the word-final sibilants from the reading task are classified as fully voiced. Complete voicelessness is more common in external sandhi (27%) than in the context of word-medial phonologically voiced sibilants (1.5%).

<table>
<thead>
<tr>
<th></th>
<th>Word-medial [+voice] dental</th>
<th>Dental sibilant in external sandhi</th>
</tr>
</thead>
<tbody>
<tr>
<td>Voiceless</td>
<td>1 (1.5%)</td>
<td>17 (27%)</td>
</tr>
<tr>
<td>Partially (de)voiced</td>
<td>7 (10.6%)</td>
<td>8 (12.7%)</td>
</tr>
<tr>
<td>Fully voiced</td>
<td>58 (87.9%)</td>
<td>38 (60.3%)</td>
</tr>
<tr>
<td>TOTALS</td>
<td>66 (100%)</td>
<td>63 (100%)</td>
</tr>
</tbody>
</table>

Table 3: Phonologically voiced dental sibilants in word-medial intervocalic position versus word-final sibilants in external sandhi (text reading): $\chi^2$ (df 2, n 129) = 18.40, p < 0.0001

Penny (1993:80-1, 2000:182) states that word-final sibilants are voiced before vowel-initial words in JS, as seen in (7). The data in Table 3 generally support this claim as a total of 73% of word-final intervocalic sibilants fall into either the fully or partially voiced category, indicating that they are voiced between one third and one hundred percent of their total duration. However, the results also suggest that external sandhi voicing should be described as a variable process, perhaps comparable to that of voicing assimilation in consonant clusters (recall (2a,b)), rather than as a completely consistent phenomenon. In fact, further evidence of variability comes from the carrier phrase items in which word-final prevocalic /s/ appears in different syntactic contexts (e.g., las amigas 'the friends', flores ermozas 'beautiful flowers', estamos aki 'we are here', diyas i semanas 'days and weeks'). As shown in Table 4, the differences observed in phonetic voicing categories are significantly dependent on the type of syntactic boundary occurring after the sibilant. In general, fully voiced realizations appear to be favored in the Det + N context (54.8%). On the other hand, voiceless realizations are most frequent in the N + Conj + N context (84.6%).

<table>
<thead>
<tr>
<th></th>
<th>Det + N</th>
<th>N + Adj</th>
<th>V + Adv</th>
<th>N + Conj + N</th>
</tr>
</thead>
<tbody>
<tr>
<td>Voiceless</td>
<td>15 (35.7%)</td>
<td>17 (42.5%)</td>
<td>22 (52.4%)</td>
<td>33 (84.6%)</td>
</tr>
<tr>
<td>Partially (de)voiced</td>
<td>4 (9.5%)</td>
<td>9 (22.5%)</td>
<td>9 (21.4%)</td>
<td>1 (2.6%)</td>
</tr>
<tr>
<td>Fully voiced</td>
<td>23 (54.8%)</td>
<td>14 (35%)</td>
<td>11 (26.2%)</td>
<td>5 (12.8%)</td>
</tr>
<tr>
<td>TOTALS</td>
<td>42 (100%)</td>
<td>40 (100%)</td>
<td>42 (100%)</td>
<td>39 (100%)</td>
</tr>
</tbody>
</table>

Table 4: Word-final prevocalic sibilants in different syntactic contexts (embedded in carrier phrases): $\chi^2$ (df 6, n 163) = 30.02, p < 0.005

4.3 Phonetic underspecification and sibilant voicing in external sandhi

What these data seem to suggest is that sibilant voicing in external sandhi is much more variable than in word-medial position, where it serves to establish a phonological contrast with voiceless sibilants. The difference between the realizations of word-medial intervocalic /z/ and word-final intervocalic sibilants can be understood in terms of the Permanent Neutralization Hypothesis (Ernestus 2003). This hypothesis posits a distinction between obstruents that are specified as either [+voice] or [–voice] in the input and output of the phonological component and others that are completely unspecified, or neutral, with regard to voicing, represented as [0voice] (also see Steriade 1997).
Phonologically voiced or voiceless obstruents require specific articulatory gestures designed to ensure that they will be perceived in accordance with their underlying voicing specification. As described above, the production of voiced obstruents always involves a reasonable degree of articulatory effort, and the realization of voiceless obstruents also necessitates specific glottal adjustments when these sounds are adjacent to sonorants. Neutral obstruents, conversely, do not have perceptual targets and do not entail any specific articulatory gestures. They are marked as [0voice] by the grammar and remain unspecified into the phonetic implementation component (hence the term ‘phonetic underspecification’). Such sounds adopt the laryngeal configurations of contiguous sounds and can therefore be expected to exhibit gradient and variable voicing as a result of the interpolation of contextual glottal activity.

The Permanent Neutralization Hypothesis is illustrated in (8), as applied to the dental sibilant voicing contrast of JS. The idealized forms in (8a,b) represent a potential underlying contrast between voiced and voiceless sibilants in word-final and preconsonantal positions, respectively. We assume that the grammar contains some mechanism depriving syllable-final sibilants of their voicing specifications, resulting in a sibilant that is neutral in [voice], represented here as [S]. (Note: In the surface forms in (8), the vertical line denotes a word boundary, and periods denote syllable boundaries.) Depending on one’s theoretical inclination, neutralization can be formalized in a number of ways. A rule-based framework might posit an autosegmental rule that delinks voicing features, while a constraint-based approach might achieve the same result by ranking a markedness constraint against contrastive voicing features in coda above faithfulness to sibilant voicing specifications in the input. The underlying contrast depicted in (8c) is unaffected by the neutralization of sibilant [voice] in coda, since both /z/ and /s/ are syllabified as onsets between vowels. In the output of the phonology, syllable-final sibilants are phonetically underspecified for voicing, while syllable-initial intervocalic sibilants maintain their underlying voicing specifications intact.

We propose that the variability of external sandhi voicing observed in Tables 3 and 4 is the result of variable and gradient interpolation between glottal targets, and that this occurs in the phonetic implementation of phonological surface forms containing neutral sibilants. For purposes of illustration, Figure 4 contrasts the implementation of a [0voice] sibilant in word-final intervocalic position (a) with that of a word-medial intervocalic [s] that is phonologically specified as [–voice] (b). We assume that resyllabification has applied in (a), moving the word-final sibilant to the onset of the following vowel-initial word. In both examples, solid horizontal lines denote glottal targets that correspond to phonologically specified [voice] features, and dotted lines show interpolation between targets. Since neutral [S] has no specified target, glottal vibration during the sibilant constriction period is determined by gradient interpolation between the surrounding vowels. As the dotted lines in (a) show, there is a range of possible trajectories that interpolation may follow. The realization of neutral [S] depends on phonetic factors such as sibilant duration and intensity, stress, adjacency to major syntactic and/or prosodic boundaries, speech register, and speaking rate. Sibilants whose duration extends beyond certain durational thresholds tend to passively devoice for aerodynamic reasons (recall Section 4.1), and voiceless fricatives are typically longer than voiced ones (Kirchner 1998:163, 236, Widdison 1997). Therefore, shorter constriction durations increase the probability of complete voicing throughout neutral [S], whereas longer durations favor gradient degrees of voicelessness. In contrast to the phonetically variable [S], the intervocalic [s] in (b) has a target for voicelessness because it is phonologically specified as [–voice]. Interpolation from the first vowel to the sibilant and from the...
sibilant to the second vowel produces transitional glottal vibration at the margins of the sibilant constriction.

Figure 4: Variable and gradient sibilant voicing as interpolation between phonetic targets

The phonetic realization of most word-medial intervocalic sibilants in our data set corresponds with the sound’s phonological voicing category, and the word-final prevocalic sibilants exhibit exactly the type of gradient voicing pattern that would be expected of [voice]-neutral obstruents. These patterns seem to fit nicely with Ernestus’ Permanent Neutralization Hypothesis and with the glottal interpolation model illustrated in Figure 4. Recall that our informant often realized phonologically voiceless sibilants with some phonetic voicing in word-medial position and that this voicing was localized in the V-C margin. This effect can be understood in the interpolation model shown in part (b) of Figure 4. Specifically, a longer transition between the glottal target of the vowel and that of the following sibilant would account for the maintenance of glottal adduction during the initial portion of the sibilant constriction. Finally, the fact that word-medial phonologically voiceless sibilants are sometimes realized as fully voiced seems to suggest an occasional failure to achieve the glottal abduction target for these sibilants, perhaps as an effect of effort minimization in intervocalic position (see the discussion in Section 4.1).

The phonetic underspecification approach can also account for the variable and gradient voicing of preconsonantal sibilants observed in modern Spanish, as shown in (2). An appeal to the [0voice] category is motivated on a purely descriptive level by the fact that most contemporary studies of preconsonantal voicing assimilation in Spanish highlight its style-dependent, gradient, and variable nature—all of which are hallmark characteristics of a phonetic process (see Harris 1969, Hooper 1972, and Navarro Tomás 1977, among others). Given a sequence such as [VS̪.ð], glottal interpolation from the preceding vowel through the [0voice] sibilant to the following voiced consonant can produce a range of trajectories, as shown for [V.S̪|V] in Figure 4. On the other hand, the failure of word-final intervocalic sibilants to undergo voicing in contemporary standard Spanish requires some rule or constraint to ensure a phonological [–voice] specification in this environment, effectively neutralizing the segmental difference between (a) and (b) in Figure 4 in favor of the latter.

5. Phonological innovation in the rhotic system of Istanbul JS

In other respects, modern JS is phonologically innovative vis-à-vis medieval Spanish and other contemporary varieties. For example, Castilian Spanish contrasts a multiple trill [r] and a single tap [ɾ] in intervocalic onsets (9a) but requires [r] in all other syllable-initial contexts, i.e., word-initially (9b) and postconsonantally (9c) (Navarro Tomás 1977).

5 See Martínez-Gil (2003) and the references cited therein for several recent analyses in Optimality Theory that treat gradient, partial voicing assimilation in Spanish obstruents as phonological. Interestingly, Martínez-Gil (2003:57) acknowledges that “[f]rom our present perspective, however, it appears that such attempts may have been premature or misconceived: I do not know of any compelling evidence suggesting that partial voicing assimilation is a phonological property, and not simply a fact of phonetic implementation. In fact, most available descriptions clearly indicate that the process is gradient, and thus typical of phonetic phenomena.”

6 The voicing of word-final intervocalic sibilants has been reported for highland Ecuadorian varieties of modern Spanish spoken around Quito and Cuenca (Lipski 1989, Robinson 1979, and Toscano Mateus 1953). Unlike JS, however, sibilant voicing is no longer contrastive in word-medial intervocalic contexts. For a complete analysis of both historical and synchronic sibilant voicing patterns in Spanish, developed in constraint-based Dispersion Theory (Flemming 1995, Padgett 2003a,b,c), see Bradley and Delforge (In press).
Penny (1992:136-7, 2000:179) claims that no variety of JS maintains a contrast between /r/ and /ɾ/. He views the merger of the contrast to /ɾ/ as an innovation not seen elsewhere in the Spanish-speaking world (except for Creole varieties such as Papiamentu and Palenquero), whereby the phonology of JS is simplified. However, others have reported maintenance of the trill (Harris 1994) and free variation between either rhotic (Malinowski 1979, Nemer 1981) depending on the dialect of JS. According to Harris (1994:76), “in Salonika, Constantinople, and Monastir there are variants of r, the single orthographic r being pronounced as a single [ɾ] and the double orthographic rr is usually a multiple [r] pronounced with two or three taps.” This description obviously contradicts Penny’s claim that all varieties of contemporary JS have neutralized the rhotic contrast.

To investigate the nature of innovation in the rhotic system of our Istanbul JS informant, we carried out a spectrographic analysis of syllable-initial rhotics produced during the same reading task employed in the investigation of sibilant voicing, as described in Section 3. Additional tokens were collected from spontaneous speech in a discussion about topics of daily life, recorded by one of the investigators. For our informant, the multiple trill [ɾ] is exceedingly rare and occurs only sporadically in emphatic speech in contexts corresponding to double orthographic <rr> between vowels, as noted by Harris (1994:76). However, we propose a different generalization about spontaneous speech based on our observations of GO’s productions: A syllable-initial rhotic in any position is a voiced approximant [ɹ] of variable duration, which often exhibits aperiodic turbulence in the upper spectra.

The representative spectrogram in Figure 5 illustrates GO’s pronunciation of the word korrer ‘to run’ in spontaneous speech. The continuous formant structure throughout the constriction period of the intervocalic rhotic indicates an approximant stricture. This realization is similar to that of the word-final rhotic, although final position shows a greater tendency toward frication, as indicated by the presence of some aperiodic energy at higher frequencies. In contrast, standard descriptions of the Spanish multiple trill in (9) note its inherently salient acoustic structure, consisting of vocalic formant values briefly interrupted by periods of stop-like silence (Blecua 2001:237-68, Widdison 1997:190). The intervocalic voiced approximant [ɹ] in Figure 5 clearly lacks this alternating acoustic structure.

The spectrogram of the phrase para komparar ‘in order to compare’, also from spontaneous speech, is shown in Figure 6. Comparing the intervocalic rhotic of para with that of komparar, we see that both show continuous formant structure, which again is indicative of an approximant constriction. Standard descriptions of the Spanish tap /ɾ/ emphasize its extremely short constriction period. Quilis (1993:337-42) observes that for Castilian Spanish, the average duration of contact is 20 ms, and that the constriction is seldom a complete closure. The first rhotic appearing in Figure 6 approximates this
duration, although here we employ the phonetic symbol [ɹ] to emphasize the approximant constriction that is typical of GO’s rhotics in JS. However, the constriction duration of the intervocalic approximant in komparar is much greater, exceeding the durational value normally associated with Castilian intervocalic /ɾ/. This durational variability is common in our corpus data and even encompasses those positions in which Castilian Spanish favors the multiple trill, as in (9b,c).

In contrast to the word-final rhotic in Figure 5, the final rhotic of komparar in Figure 6 exhibits greater aperiodic noise in the upper spectra as well as the complete absence of formant structure or voicing. This difference is plausibly attributed to the fact that the word-final rhotic in Figure 6 is also phrase-final. It is interesting to note that in many varieties of Latin American Spanish, prepausal rhotics are often realized as strident fricatives, either with or without devoicing. (See, for example, the comprehensive dialectal descriptions of Lipski 1994.) One possibility, therefore, is that the longer, voiceless fricative realization exemplified in Figure 6 reflects a universal phonetic tendency toward phrase-final devoicing and/or lengthening. What seems more likely, given the discussion of GO’s multilingual background in Section 3, is that this realization is the result of phonological transfer from Turkish, which has a process of final liquid devoicing. Kornfilt (1997:486-7) notes that in the variety of Turkish spoken in Istanbul, the liquids /ɾ/ and /l/ are realized as voiced between vowels in (10a) but as voiceless word-finally before pause in (10b).

(10) a. kare  kare  'square'
    hala  hala  'father’s sister'
 b. kʲaɾ̥ kär  'profit'
    kʲel  kel  'bald'

Moreover, Kornfilt observes that in Istanbul Turkish, devoicing of the tap occurs with greater frequency, and female speakers tend to devoice both liquids more often than males do. We observe that these patterns are reflected in the JS speech of our informant, who occasionally pronounces prepausal lateral liquids as devoiced with some frication. This type of realization is illustrated by the spectrogram in Figure 7, taken from our spontaneous speech recordings.
Our descriptive results for Istanbul JS rhotics are in agreement with Hammond’s (1999, 2000) findings for contemporary dialects of Latin American Spanish, namely that the prescriptive multiple trill simply does not occur in normal discourse of the vast majority of speakers. Cross-linguistically, trills are known to alternate with rhotics that lack complete occlusions, such as fricatives or approximants (Blecua 2001, Colantoni 2001, Lewis 2004, Lindau 1985, Widdison 1997). With respect to Penny’s (1992, 2000) claim that all contemporary varieties of JS have neutralized the contrast between /r/ and /ɾ/ in favor of the latter segment, the data from our Istanbul JS informant are suggestive of a system in which the prototypical rhotic realization in syllable-initial position is a voiced approximant [ɹ] of variable duration. However, we do not wish to claim that there is neutralization of the intervocalic contrast. While the multiple trill [r] is absent from GO’s spontaneous speech productions, it is conceivable that rhotic contrast might be maintained via other acoustic cues that are transitional or internal to the approximant constriction (e.g., differences in duration and/or degree of stridency, formant transitions in the C-V or V-C margins, coloring and/or duration of the preceding vowel, etc.). Further investigation is necessary to determine whether such cues exist and whether listeners are able to perceive a rhotic contrast by attending to such cues.

6. Conclusion

In this paper, we have presented a phonological and phonetic description of sibilants and rhotics in Istanbul JS, based on the acoustic analysis of recordings of an elderly female speaker of this variety. To avoid potential complications associated with the comparison of JS speakers from highly multilingual backgrounds, we collected and analyzed acoustic data from a single informant speaking in different tasks. In general, our findings suggest retention of the medieval Spanish contrast between voiced and voiceless sibilants in word-medial intervocalic position, while the occurrence of phonetic sibilant voicing in external sandhi contexts is much more variable than previous descriptions of JS have indicated. This difference is adequately captured by a phonetic underspecification approach to voicing neutralization, which was shown to account for the gradient and variable nature of sibilant voicing.

Based on an acoustic analysis of rhotics in syllable-initial contexts, we observe that our informant generally favors an approximant realization to the exclusion of the multiple trill that appears in modern Castilian, although this does not necessarily entail the neutralization of rhotic contrast. Therefore, the innovative behavior of Istanbul JS rhotics fits in with many other contemporary Spanish varieties in which the multiple trill is largely absent from normal discourse. However, there is also evidence of local innovation due to phonological transfer from Turkish final liquid devoicing, which seems to have influenced the realization of final liquids in our informant’s JS speech.

In sum, we have contributed novel empirical data and generalizations from the variety of JS spoken in Istanbul. Given the endangered status of modern JS, we encourage further investigation of the phonological and phonetic variation observed across extant varieties of this language.
References


Padgett, Jaye. 2003c. Systemic contrast and Catalan rhotics. Santa Cruz: University of California at Santa Cruz, ms.


