Spanish is a Romance language spoken by approximately 405,638,110 speakers in the world (Lewis, Simons & Fenning 2013). Two major varieties are distinguished, Peninsular Spanish (Spain) and the Spanish spoken in the Americas, although it is also spoken natively in some parts of Africa, and in the United States. Spanish in the Americas comprises several dialects well differentiated by variations in the lexicon, phonology and, more importantly, in intonational patterns. In Mexico 86,211,000 (88% of the population) use Spanish as their first language, and a significant number of indigenous people have Spanish as their second language. The variety illustrated here is representative of the speech of the educated middle-class population from the metropolitan zone (three female and three male speakers in their 30s), which has as its center Mexico City, the most densely populated urban area in the country with more than 20 million people according to the Mexican National Census (INEGI 2010).

Consonants
The consonant inventory of Mexico City Spanish is presented in the chart below. The consonants in parentheses occur primarily in loanwords.

<table>
<thead>
<tr>
<th>Consonant Type</th>
<th>Bilabial</th>
<th>Labio-</th>
<th>Denti-</th>
<th>Post-</th>
<th>Velar</th>
</tr>
</thead>
<tbody>
<tr>
<td>Plosive</td>
<td>p b</td>
<td>t d</td>
<td></td>
<td>k g</td>
<td></td>
</tr>
<tr>
<td>Affricate</td>
<td></td>
<td></td>
<td>(tʃ)</td>
<td>(tʃ)</td>
<td></td>
</tr>
<tr>
<td>Nasal</td>
<td>m n</td>
<td></td>
<td></td>
<td>n</td>
<td></td>
</tr>
<tr>
<td>Fricative</td>
<td>f s</td>
<td>(ʃ) j</td>
<td>(x)</td>
<td></td>
<td></td>
</tr>
<tr>
<td>Lateral</td>
<td>l</td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Flap</td>
<td>r</td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Trill</td>
<td>r</td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Approximant</td>
<td></td>
<td>j w</td>
<td></td>
<td></td>
<td></td>
</tr>
</tbody>
</table>
Plosives /t/ and /d/ are typically dento-alveolar whereas the rest of the coronals, /n s l r t/, are generally alveolar. Voiceless plosives are characterized by a very short VOT: /p/ = 15 ms, /t/ = 20 ms, /k/ = 49 ms, whereas voiced plosives have rather long negative VOT values: /b/ = −65 ms, /d/ = −84 ms, /ɡ/ = −73 ms (mean values from three females, and three males speakers, n = 100 tokens per consonant). Glottal stop is non-phonemic in Spanish but it can be inserted as onset to onsetsyllable (as in e.g. [ʔel ña] ‘the traveler’ in the transcribed passage below). As in other varieties of Spanish, in Mexico City Spanish one of the mainallophonic variations of the voiced series is lenition in intervocalicenvironments; other contexts in which lenition can occur include word-final position or after a vowel and before a liquid. In Mexico City Spanish the allophones range from fricatives to approximants when preceded by /t/ or /d/. In Mexico City Spanish the allophones range from fricatives to approximants when preceded by /r/. In Mexico City Spanish the allophones range from fricatives to approximants when preceded by /n/ or /ɲ/. Clusters involving nasals as the first member are always homorganic, [imˈpoːsībl] ‘impossible’, [ˈmango] ‘mangofruit’, [imˈfleliʃ] ‘unhappy’, [ˈiŋjekʃon] ‘injection’. Place assimilation may occur across words provided that no pause is made in between the nasal and the following consonant (see [tamˈfwerter] ‘as strong as’ in the transcribed passage below). The palatal nasal occurs mainly in intervocalicposition. However, some innovative words of Mexico City Spanish, such as [ˈnaŋara] ‘tickle’, [ˈnɛɾo] ‘buddy’ or [ˈnoŋo] ‘silly’, show that the phonotactics of the language allows word-initial palatal nasals.

The lateral in Mexico City Spanish is alveolar, but varies its place of articulation to dental or alveolar according to the following segment [el] ‘from the north’ and [emˈbweɾto] ‘wrapped’. The lateral is becoming rather unstable in word-final position: it can vary between devoiced, unreleased, reduced or deleted (e.g. ‘cual de los dos’ [ˈkwa de los ˈdos] in the transcribed passage below).

Spanish has a phonemic contrast between a flap and a trill, /ɾ/ ~ /ɾ/. The distribution of this pair is defective in the sense that the flap occurs intervocally, word-finally, and in consonant clusters (plosive–flap), whereas the trill occurs word-initially and intervocally (see Harris 1969 for phonological analysis). There is a previously undescribed realization of the trill in the speech of some speakers of Mexico City Spanish. The most salient observation indicates that the trill is also realized as a voiced retroflex fricative [z], especially noticeable in the speech of middle-class female speakers. Such pronunciation was also found in the speech of some the speakers included in this illustration. Figure 1 includes the sound pressure oscillograms and corresponding spectrograms illustrating the variation of the phonemic trill in two female speakers (f2 and f3).

One of the main differences between Peninsular Spanish and the Spanish spoken in the Americas is the absence of [θ] in the latter. The corresponding sound is pronounced as [s] in Mexico City Spanish. Frequently, [s] is fully or, more often, partially voiced before voiced
Figure 1 Spectrograms and waveform of words illustrating the realization of /r/ as [r] and [z] in the word perro ‘dog’ in the speech of two female speakers (f2) and (f3). The left panel shows the trill and the right panel shows the fricative pronunciation.

consonants [‘izla] ‘island’, [‘azma] ‘asthma’, [‘sizne] ‘swan’, [izra’el] ‘Israel’, [ez’boso] ‘sketch’, [‘dezde] ‘from’, [muzyo] ‘mold’. Mexico City Spanish has a velar fricative [x] rather than glottal or uvular [h] or [ɬ], which occur in other varieties. The production of the palatal fricative /ʝ/ varies in a continuum from an approximant to a fricative to an affricate, [j] ~ [ʝ] ~ [ɭ]. The variation depends on individual, stylistic and even lexical factors such as a word frequency. The approximant is frequently found in fast speech, as well as in some words written with hi, like in hiedra ‘ivy’, hiel ‘bile’, hierro ‘iron’. However, some words like yelmo ‘helmet’ and yerro ‘I wander (accidentally from a fixed or chosen route)’ are pronounced mainly as affricates. The affricate allophone tends to occur word-initially, as in [ʃeno] ‘full’, and after sonorants, as in [kɔνjuʃe] ‘spouse’. The fricative and the approximant are equally common between vowels [kaje] ~ [kæj] ‘street’.

Influence from American Indian languages
Mexico City Spanish has been influenced by contact with native languages of the Americas for more than five centuries. In the case of the variety spoken in Mexico City the most notable influence comes from Nahuatl, the language spoken by the Aztecs, and still the indigenous language with most native speakers in Mexico. The Nahuatl substrate introduced sounds such as the voiceless palatal fricative [ʃ], the voiceless alveolar affricate [ts] and the voiceless alveolar lateral affricate [ɭ] into Mexico City Spanish. Illustrations of these sounds are shown below with spectrograms of [ʃola] ‘Xola (name of a street and neighborhood in Mexico City)’, [askapo’tsalko] ‘Azcapotzalco (name of a neighborhood)’ and [popoka’tepetl] ‘Popocatepetl (name of a volcano)’, in Figures 2, 3 and 4, respectively. As a result of this influence, words with a lateral affricate are quite common in Mexico City Spanish, while in other Spanish dialects they are limited to a few words or completely absent. Yet, the adaptation into Mexico City Spanish loan words containing [ʃ] and [ɭ] is subject to a wide range of variation, while [ts] is rather stable. Thus, for instance, [ʃ] is also pronounced as [ɭʃ], [ɭɭ] occurs mainly in coda position, while a [ɭɭ] one occurs mainly in onset position.
Figure 2  Illustration of the borrowed sound /ʃ/ from Nahuatl in Mexico City Spanish with the word ['ʃola].

Figure 3  Illustration of the borrowed sound /ts/ from Nahuatl into Mexico City Spanish with the word [askapo'tsalko].

Figure 4  Illustration of the borrowed sound /tl/ from Nahuatl in Mexico City Spanish with the word [popoka'tepeh].
Spanish has a five-vowel system /a e i o u/. The distribution of the vowels in the acoustic space is typical of a five-vowel system, with the front vowels having a higher position, i.e. lower values of F1. Figure 5 shows the mean values of female and male vowels. The five vowels can appear in both stressed and unstressed positions; /i/ and /u/ can be realized as close vowels in unstressed positions. Overall, stressed vowels are typically longer than unstressed ones. There is a tendency to use more open pronunciations of vowels in closed syllables.

Diphthongs are formed with [i̯] or [u̯] as offglides in combination with any of the other vowels (excepting for the sequences u̯, i̯) /aí aú eí eu oí/. In Mexico City Spanish, as in other varieties (Chitoran & Hualde 2007), some words show alternation between a diphthongal pronunciation and vowel hiatus, e.g. [ˈpijano] ‘piano’ versus [piˈano] ‘piano’. However, there are lexical restrictions to the alternation, as pairs of words with diphtong or hiatus may contrast as in hoy ‘today’ and oí ‘I heard’. In running speech, diphthongs may also be formed with a mid glide, [e̯] or [o̯], as in oleo ‘oil painting’ or poeta ‘poet’.

Final lenition, expressed as the reduction and even deletion of unstressed vowels in syllables final within the scope of an intonational phrase, is typical of Mexico City Spanish. Words like [ˈfwertʃ] ‘strong’, [poðeˈrosɔ] ‘powerful’, [iŋtensə,menʧ] (spelled out as intensamente) ‘intensely’ are examples observed in the transcribed passage below. The following examples illustrate the typical variation in lenition ranging from vowel shortening to vowel devoicing and vowel deletion as the most extreme case:

\[
\begin{array}{ll}
\text{a} & \text{‘aso aso} \text{ ‘I roast’} \\
\text{e} & \text{‘seso seso} \text{ ‘brain’} \\
\text{i} & \text{‘išo hizo} \text{ ‘He made’} \\
\text{o} & \text{‘oso oso} \text{ ‘bear’} \\
\text{u} & \text{‘uso uso} \text{ ‘use’} \\
\end{array}
\]

\[
\begin{array}{ll}
\text{ai} & \text{‘ai hay} \text{ ‘there is’} \\
\text{au} & \text{‘au causa} \text{ ‘cause’} \\
\text{ei} & \text{‘ei ley} \text{ ‘law’} \\
\text{eu} & \text{‘eu deuda} \text{ ‘debt’} \\
\text{oi} & \text{‘oi muy} \text{ ‘very’} \\
\end{array}
\]

\[
\begin{array}{ll}
\text{aj} & \text{‘aj hay} \text{ ‘there is’} \\
\text{aș} & \text{‘aș causa} \text{ ‘cause’} \\
\text{eț} & \text{‘eț ley} \text{ ‘law’} \\
\text{euț} & \text{‘euț deuda} \text{ ‘debt’} \\
\text{oț} & \text{‘oț hoy} \text{ ‘today’} \\
\text{uț} & \text{‘uț muy} \text{ ‘very’} \\
\end{array}
\]

\[
\begin{array}{ll}
\text{aj} & \text{‘aj hay} \text{ ‘there is’} \\
\text{aș} & \text{‘aș causa} \text{ ‘cause’} \\
\text{eț} & \text{‘eț ley} \text{ ‘law’} \\
\text{euț} & \text{‘euț deuda} \text{ ‘debt’} \\
\text{oț} & \text{‘oț hoy} \text{ ‘today’} \\
\text{uț} & \text{‘uț muy} \text{ ‘very’} \\
\end{array}
\]
Stress and intonation

Spanish stress is contrastive and can fall on any of the three last three syllables of a word. Penultimate stress is the most frequent pattern; however, stress occurs in the pre-antepenultimate syllable in verbs containing proclitic forms (such as in imperative forms) [reˈyaləˌselo] ‘give it to her!’.

Figure 6 shows the spectrogram, intensity curve, f0 contour and sound pressure oscillogram of a triple contrast of stress. As the figure shows, there is no single acoustic cue that can signal stress alone, but the prominence of the stressed syllable is achieved by the three acoustic properties together (see Ortega-Llebaria & Prieto 2011 for stress in Catalan and Castillian Spanish.)

There are two major intonational patterns of declarative sentences in Mexico City Spanish. One is similar to that of varieties described in earlier studies (Navarro Tomás 1944, De la Mota et al. 2010) in that the f0 peak of a pitch accent is shifted from the stressed syllable to the following syllables. The other intonational pattern is emblematic of the working class speech variety of Mexico City Spanish, where the peak of the pitch accent is aligned with the stressed syllable (see Prieto, van Santen & Hirschberg 1995, Kim & Avelino 2003).

Transcribed passage

Orthographic version

El viento del norte y el sol estaban discutiendo sobre cuál de los dos era el más fuerte, cuando pasó un viajero envuelto en una gruesa capa. Quedaron de acuerdo en que quien primero lograra que el viajero se quitara la capa sería considerado el más poderoso. Entonces, el viento del norte sopló tan fuerte como pudo, pero entre más soplaba, más se cobijaba en su capa el viajero; al fin el viento del norte se rindió. Entonces el sol brilló intensamente, e inmediatamente el viajero se quitó la capa. Así, el viento del norte se vio obligado a reconocer que el sol era el más poderoso de los dos.
Figure 6 Spectrogram, intensity curve, pitch track, and sound pressure wave form of the triplet [ˈselebɾe] 'celebrated', [seˈlebre] 'he celebrates (SUBJUNCTIVE)', and [seleˈbre] 'I celebrated' illustrating the stress contrast in Mexico City Spanish.

Narrow transcription
el ˈbjen̞to ðel ˈnorte || jelˈsol || ćs'taβan diskv'tjeŋdo soβre ˈkwa de los ˈdos || era əl ˈmas ˈfwertɛ || kwando ˈpaʃo ŋmbjaˈxeremˈbwelo enˈuna ˈgrwesa ˈkapʰa || keˈðaron de aˈkwerdo || eŋke ˈkjem ˈprimero loˈyrrara kel βjaˈxero se kˈtara ˈla ˈkapa || seˈria konsiðeˈraðo el ˈmas poðeˈrosɔ || enˈtonses || el ˈbjen̞to ðel ˈnorte sɔˈplo tɑŋ ˈfwerte komo ˈpuðo || ˈpərə entre masoˈplaba || ˈmase kɔβiˈxaβaŋ su ˈkapaŋ βjaˈxero || ʔal ˈtʃin ʔel ˈbjen̞to ðel ˈnorte se riɲˈdjo || enˈtonses el ˈsol bɾiˈjo ʃtʃenˈtensəmɛnt || ʔeɪməˈdʒatɛmɛnt ʔel βjaˈxero se kˈtɔ la ˈkapə || ʔaˈsi|| el ˈbjen̞to ðel ˈnorte || seˈβjobli ɣədaˈɾəkonoˈseʃə || kel ˈsol || əɾəɾ ˈmas poðeˈrosø de los ˈdɔs

Acknowledgements
I am grateful to Juana Gil Fernández, Ascensión Hernández Triviño y Claudia Parodi for their comments and suggestions. Special thanks to Adrian Simpson, Amala Arvaniti and Ewa Jaworska for editorial guidance. I acknowledge the anonymous reviewers for their opportune comments. Finally, many thanks to the speakers who volunteered to record their voices for this illustration. I am solely responsible for the contents of this article.
This research was carried out while the author was at the Department of Linguistics, University of California, Berkeley.

References


