History of Spanish Language Irregularities

How does knowing the history of Spanish help us understand some of the puzzling structures or "irregularities" found in the modern language? In order to answer this question, one must understand the nature of the different types of linguistic analysis, of which there are basically two: **synchronic** and **diachronic**.

The nature of **synchronic analysis**, i.e., analysis of the language at one given point in time (usually the modern language), is such that it describes, i.e., it provides a description or statement of the facts that we possess in the modern language. **Diachronic or historical analysis**, on the other hand, often explains how these facts came to be as they are.

Let us look at one example. Synchronously, we can state that the indirect object pronouns le/les become or are replaced by se when appearing before the direct object pronouns lo, la, los, las. Without looking historically, we might try to account for this phenomenon with unfounded, ad hoc explanations like "le lo sounds bad", or "lelo means 'stupid'". Historically, however, we find that this phenomenon resulted from a series of normal or expected phonological (sound) changes: Latin ILLI ILLUM > (IL)LI ILLUM > li ellu > ljelo > yelo > gelo (the g pronounced like j as in John) > šelo (the initial consonant pronounced like the sh of English) > selo (the last step by analogy to reflexive se and confusion of š (like sh) and the apical or Castilian s).

Thus, upon becoming acquainted with the history of the language, one comes to realize that many of the anomalous structures or so-called "irregularities" of Spanish, from a historical perspective, are not really irregular at all, but rather, are the result of a progression of regular changes from Latin to Spanish.

What follows is a list of common questions about some of the "irregularities" of Spanish, accompanied by answers to these questions.

**QUESTION:** Why do some verbs have diphthongs -ie or -ue in some forms like the singular:

ILLAS AQUAS > ilas aguas > elas aguas > (e)las aguas > las aguas

Thus, if students realize that historically this is not the masculine definite article, they will not confuse the gender of the noun when modifying it, i.e., el agua pura, not *el agua puro (unlike el poema épico, un poeta famoso, etc.). (*incorrect form)

**QUESTION:** Why do the pronominal forms conmigo, contigo, consigo have a -go on the end?

The way to say "with me", "with you", "with him/her self, themselves", in Latin was MECUM, TECUM, SECUM, which were really a combination of the object pronouns ME, TE, SE plus the preposition CUM, which was commonly attached to the end (thus, a postposition, rather than a preposition). Through regular sound change, these forms evolved as follows:

MECUM > mec > meco > mega
TECUM > tec > teco > tego
SECUM > seco > seco > sego

(Compare Latin AMICUM > amicu > amico > amigo)

Once the forms had reached this stage, the -go element, originally CUM, was beginning to become unrecognizable and lose its meaning, since CUM everywhere else had evolved to con. So speakers added the preposition con to the beginning of these forms, producing conmego, contego, consego, which quickly shifted to conmigo, contigo, consigo because the pronouns mi, ti, si (not me, te, se) followed all other prepositions.

**QUESTION:** Why do some verbs have diphthongs -ie or -ue in some forms but not in others?

Latin had two types of /e/ and two types of /o/. One set was more close, like the stressed vowels in peso and todo, while the other was more open, like the stressed vowels in perro and torre. The more open vowels, when carrying vocal stress, underwent a process of diphthongization, as follows:

\[
\begin{align*}
\text{é} & \rightarrow \text{ée} \\
\text{ió} & \rightarrow \text{úo} \\
\text{Í} & \rightarrow \text{úe} \\
\end{align*}
\]
When not stressed, these vowels did not diphthongize. A verb like *negar*, then, will have –*ie*– only in the stressed syllable, thus: *niego*, *niegas*, *niega*, *niegan*; but, *negamos*, *negáis*, *negado*, etc. Likewise, in a verb like *poder* we find *puedo*, *puedes*, *puede*, *pueden*; but, *podemos*, *podéis*, *podido*, etc. Verbs that had the more close variety of /e/ and /o/ do not show diphthongs in any forms, thus: *debo*, *debes*, *debe*, *debemos*, *debéis*, *deben*; *como*, *comes*, *come*, *comedos*, *coméis*, *comen*.

**QUESTION:** Why, then, do verbs *tener* and *venir* not have the diphthong –*ie*– in the first person singular, i.e., *tengo*, *vengo*?

The /e/ in these forms should have diphthongized, but did not, because of an element in the suffix known as the "yod". This element had the effect of raising the more open /e/ to the more close type found in verbs like *debo*, *debes*, etc., thus impeding diphthongization, as follows:

TENO > *tenjo* (containing open /e/ and yod [j]) > *tenjo* (containing close /e/ and yod).

VENEO > *venjo* (containing open /e/ and yod [j]) > *venjo* (containing close /e/ and yod).

The yod in the forms *tenjo* and *venjo* were then replaced by /g/ (by analogy with verbs like *digo* and *hago*), yielding *tengo* and *vengo*.

Since the other forms of these verbs did not have a yod in their suffix, the stem vowels either diphthongized or not, according to whether or not they were stressed, for example:

TENES > *ténés* > *tienes*
But, TENEMUS > *tenemos*

VENIS > *vénes* > *vienes*
But, VENIMUS > *venimos*

**QUESTION:** Why do some verbs change -e- ->-i- and -o- -> -u- in only some of their forms?

Again, it is because of either the presence or absence of the yod. Just as the yod raised open /e/ and /o/ to close /e/ and /o/, it also raised close /e/ and /o/ to /i/ and /u/ respectively. In verbs like *servir* and *dormir*, one finds a yod only in the third person singular and plural of the preterite, and in the present participle, i.e., in the endings –*ió*-, –*ieron*-, –*iendo* (phonetically [jó], [jéron], [jéndo]). Thus Old Spanish *servió*, *servieron*, *serviendo* and, *dormió*, *dormieron*, *dormiendo* all shifted to modern Spanish *sirvió*, *sirvieron*, *sirviendo* and *durmieron*, *durmieron*, *durmendo*.

One may wonder why the present subjunctive forms *sirvamos*, *sirváis* and *durmió*, *durmieron*, *durmendo* show the /e/ > /i/ and /o/ > /u/ change when there is no yod in the suffixes -*amos* and -*áis*. The answer is that there was a yod in the Latin forms, which was eliminated from the suffix after it had raised the vowels, thus: SERVIAMUS > *serviamos* > *sirviamos* > *sirvamos*; DORMIAMUS > *durmiamos* > *durmamos* > *durmamos*.

It should be noted that the effect of the yod can also be found in individual words. In the word *diciembre*, for example, the yod affected the vowel of the previous syllable, raising the /e/ of the Old Spanish word to /i/, e.g., Old Spanish *deciembre* > modern Spanish *diciembre*.

**QUESTION:** Why do *ser* and *ir* have the same form in the preterite?

First, it is really the perfect (= Spanish preterite) conjugation of the Latin verb *ESSE* "to be" which came to be used for both *ser* and *ir* in Spanish. In Spoken Latin, the preposition IN (> Spanish *en*) used with a location came to mean "movement toward". Also, *ser* and *estar* did not have the same restrictions in Old Spanish that they do in modern Spanish and consequently *ser* occurred at times where one might expect *estar*. Thus FUERUNT IN CAMPUM, for example, originally meant "they were at the countryside", but later "they were moving/in route toward the countryside", and eventually came to mean "they went to the countryside". Therefore, this paradigm, *fui*, *fuiste*, etc., eventually became associated with the infinitive *ir* "to go".

**QUESTION:** Why do some verbs change -e- ->-i- and -o- -> -u- in only some of their forms?

Again, it is because of either the presence or absence of the yod. Just as the yod raised open /e/ and /o/ to close /e/ and /o/, it also raised close /e/ and /o/ to /i/ and /u/ respectively. In verbs like *servir* and *dormir*, one finds a yod only in the third person singular and plural of the preterite, and in the present participle, i.e., in the endings –*ió*-, –*ieron*-, –*iendo* (phonetically [jó], [jéron], [jéndo]). Thus Old Spanish *servió*, *servieron*, *serviendo* and, *dormió*, *dormieron*, *dormiendo* all shifted to modern Spanish *sirvió*, *sirvieron*, *sirviendo* and *durmieron*, *durmieron*, *durmindo*.

HABUI > *abui* > *aubi* > *ovi* > *ove*

The Latin perfect of *TENERE* was *TENUI*, which may have begun to evolve through regular sound change as shown below, but was eventually discontinued:

TENUI > *teuni* > discontinued
Because of the semantic similarity between HABERE and TENERE, the replacement for TENU/teuni was created as follows:

\[
\begin{align*}
\text{aver} & : \text{tener} \\
\text{ove} & : X = \text{tove} \text{ (replacing *teuni)}
\end{align*}
\]

The Latin perfect of STARE was STETI, which evolved as follows:

\[
\begin{align*}
\text{STETI} & > \text{isteti} > \text{estiti} > \text{estidi} > \text{estide} \\
\text{Old Spanish estide} & \text{, however, was swept up in the analogy shown above, as depicted below, thus producing Old Spanish estove:}
\end{align*}
\]

\[
\begin{align*}
\text{aver} & \quad \text{tener} & \quad \text{estar} \\
\text{ove} & \quad X = \text{tove} & \quad X = \text{estove} \text{ (replacing estide)}
\end{align*}
\]

The other "irregular" preterites evolved through regular sound change from their Latin ancestors as follows:

\[
\begin{align*}
\text{SAPUI} & > \text{saupi} > \text{Old Spanish sope} \\
\text{CAPUI} & > \text{caupi} > \text{Old Spanish cope} \\
\text{POTUI} & > \text{pouti} > \text{pudi} > \text{Old Spanish pude} \\
\text{POSUI} & > \text{pousi} > \text{pusi} > \text{Old Spanish puse}
\end{align*}
\]

Note that of all of these Old Spanish preterites, only those of poder and poner had the vowel \(u\) in their stem. That is, pude (and perhaps puse) caused Old Spanish ove, tove, estove, sope, and cope to shift to uve (now spelled hube), tuve, estuve, supe, and cupe.

**QUESTION:** Why is veía, and not *via, the imperfect form of ver?  

Because in Old Spanish the verb was veer, like leer and creer. Thus we have in modern Spanish veía, like leía and creía. Old Spanish veer later contracted to ver, while leer and creer have not.

**QUESTION:** Why are era, eras, etc. the imperfect forms of ser, and iba, ibas, etc. the imperfect forms of ir?  

Because both sets of forms continue the Latin paradigms with very little, though regular, sound change, as shown below:

**ESSE "to be"**

<table>
<thead>
<tr>
<th>ser</th>
<th>era</th>
<th>éramos</th>
</tr>
</thead>
<tbody>
<tr>
<td>ERAM</td>
<td>ERAMUS</td>
<td>era</td>
</tr>
<tr>
<td>ERAS</td>
<td>ERATIS</td>
<td>eras</td>
</tr>
<tr>
<td>ERAT</td>
<td>ERANT</td>
<td>era</td>
</tr>
</tbody>
</table>

**IRE "to go"**

<table>
<thead>
<tr>
<th>iba</th>
<th>ibamos</th>
</tr>
</thead>
<tbody>
<tr>
<td>IBAM</td>
<td>IBAMUS</td>
</tr>
<tr>
<td>IBAS</td>
<td>IBATIS</td>
</tr>
<tr>
<td>IBAT</td>
<td>IBANT</td>
</tr>
</tbody>
</table>

**QUESTION:** Why do the present tense forms of ir look like forms of the -ar verbs?

The present tense of this verb is actually made up of forms from two different Latin verbs, namely, IRE "to go", and VADERE "to go quickly". Some of the forms of IRE became, through regular sound change, too short, and were replaced by forms of VADERE. The forms of VADERE should have actually evolved like those of CADERE (Spanish caer "to fall"), yielding "vaes, "vae, etc., but did not. Instead, because of analogy with forms of the verb DARE "to give", Spanish dar, we find in the earliest Old Spanish texts, the forms vo, vas, va, and van (like Old Spanish do, das, da, dan from Latin DO, DAS, DAT, DANT). In fact, in Old Spanish, the entire paradigm was: vo, vas, va, imos, ides, van, showing two other forms originally from IRE, namely, imos < Latin IMUS, and ides < Latin ITIS. These were eventually replaced by vamos, vais (in the mid-15th century). The only other change from there was Old Spanish vo > modern Spanish voy, explained below.

**QUESTION:** Why do the verbs voy, soy, doy, and estoy end in -y?

These forms descended from Latin SUM, VADO, DO, STO, which, through regular sound change, evolved to so, do, estó in Old Spanish. The syntactic combinations of these verbs with the postposed subject pronoun, i.e., so yo, do yo, vo yo, estoy yo, eventually produced articulations such as soy yo, doy yo, voy yo, estoy yo, while they remained unchanged if the pronoun came before or did not appear at all. After many generations, however, the verbs themselves, that is, even without the postposed subject pronoun, were reinterpreted by speakers as soy, doy, voy, estoy. The reason why this did not occur with other verbs, for example, digo yo, creo yo, etc., is because the final -o of these is unstressed, whereas the final -o of Old Spanish so, vo, do, estó was indeed stressed. It was the stressed
nature of the final vowel of these forms that attracted an element of the pronominal form yo to them.

**QUESTION:** Why do all three conjugations (-ar, -er, -ir) have the same endings in the future? (This phenomenon applies to Portuguese, French, and Italian as well.)

Because, historically, the future forms are really a combination of the infinitive + the verb haber (Old Spanish aver). Haber was often used in place of, or where one would expect in modern Spanish, tener. If one said he/she had to do something, this obligation carried an implied futurity. There was then a shift in meaning: hablar (h)e con él "I have to speak with him" > "I will speak with him".

**QUESTION:** Why do some verbs, like morir, poner, escribir, etc., have irregular past participles?

Because these were "irregular" in Latin. For example, the forerunner of Spanish escrito was Latin SCRIPTUM, which evolved regularly as follows:

SCRIPTUM > iscriptu > escrito > escrito

Note the English derivatives, scripture, script, manuscript (manu "hand" + script "written").

We provide below a list of the most common "irregular" past participles of Spanish along with their Latin ancestors and English derivatives:

<table>
<thead>
<tr>
<th>Infinitive</th>
<th>Spanish</th>
<th>Latin</th>
<th>English derivatives</th>
</tr>
</thead>
<tbody>
<tr>
<td>Morir</td>
<td>muerto</td>
<td>MORTU(U)M</td>
<td>mortuary, mortician</td>
</tr>
<tr>
<td>Poner</td>
<td>puesto</td>
<td>POS(I)TUM</td>
<td>position, posit, post</td>
</tr>
<tr>
<td>Ver</td>
<td>visto</td>
<td>VIS(T)UM</td>
<td>vision, visor, visual</td>
</tr>
<tr>
<td>Escribir</td>
<td>escrito</td>
<td>SCRIPTUM</td>
<td>scripture, (manu)script</td>
</tr>
<tr>
<td>Abrir</td>
<td>abierto</td>
<td>APERTUM</td>
<td>aperture, apertural</td>
</tr>
<tr>
<td>Romper</td>
<td>roto</td>
<td>RUPTUM</td>
<td>rupture, erupt</td>
</tr>
</tbody>
</table>

Through the Latin connection and association with the English derivatives, students may more readily remember the -t- or -st- element of the past participle, thus avoiding the common errors *morido, *ponido, *escrito, etc.

**QUESTION:** Why does Spanish have the letter h if it is never pronounced?

There are three reasons.

First, Latin had some words that began with h-, and this letter was pronounced as it is in English. It soon ceased to be pronounced, so that the Latin words beginning with this letter appear in the oldest Spanish texts without it. For example, Latin HABERE "to have" > Old Spanish aver. The h- was eventually replaced to make these Spanish words look more like Latin in the hope that Spanish would attain the same level of prestige. Thus, Old Spanish aver, for example, was given a new spelling, haber (note that the b, which has changed to v in some texts and u in others, was also replaced.).

Second, many Latin words that began with t- underwent a change from Old to modern Spanish in which this sound came to be pronounced like h, and eventually, like the original Latin H, not at all. For example, FABULARE "to tell fables" > Old Spanish fablar "to speak" > Late Medieval Spanish hablar (in which the h was pronounced like the h of English) > modern Spanish hablar (in which the h is no longer pronounced). The now silent h- has simply been retained.

Third, in Old Spanish, the letter u could represent either a vocalic or consonantal element. Thus, the Old Spanish spelling, ueuo, could have represented either the word for "egg", or "I drink". In order to indicate that the word-initial u- was a vowel, scribes adopted the practice of adding a silent h- (which was never part of the Latin spelling). Thus, ueuo, when meaning "egg" (< Latin OVUM), was later written hueuo to indicate the vocalic nature of the u-, and eventually, huebo. The word ueuo, when meaning "I drink", was eventually respelled bebo, again, to give it a more Latin look, cf. Latin BIBO "I drink". Another example is Latin OSSUM "bone" > Old Spanish ueso (which could also be pronounced like beso "kiss") > hueso.

One notes that French also employed h- for the same reason: Latin OCTO "eight" > Old French uite (which was also the old spelling for modern French vite "quickly") > modern French huite "eight".

By Suzanne M. Varner, St. Christopher's School, Richmond, Virginia

Source:
http://www.virginia.edu/cla/_fpclass/ADVProj/ADVProj/VarnerIntro.htm
http://www.virginia.edu/cla/_fpclass/ADVProj/ADVProj/varnerhistory.html