Quasi-Phonemic Contrasts in Spanish

José Ignacio Hualde

University of Illinois at Urbana-Champaign

1. Introduction

The striking thing about phonology is that the infinite phonetic variety in the utterances of any language can be reduced to a small inventory of contrastive units or phonemes. The bad news is that oftentimes phonemization is problematic in some corners of the language. Typically if the same language has been investigated by two linguists, we will get slightly different phonemic inventories; and these differences of opinion usually do not go away as the language is studied more extensively. Both the considerable extent to which we normally find agreement among linguists in the phonemic analysis of a given language and the existence of areas of disagreement are remarkable facts.

In this paper I consider in detail three specific aspects of the phonemic inventory of Spanish that remain controversial and argue that in all three cases it is useful to speak of quasi-phonemic contrasts (although not exactly in the same sense in every case). The paper ends with some general considerations about the nature of phonological categorization.

Let me start with an anecdote. I recently submitted a paper reporting on experimental data on Basque suprasegmentals to a journal. A reader made the sensible suggestion that the examples, which were in conventional Basque orthography, should also be given in IPA transcription. This seemed reasonable enough, and I replied that I would comply. However, it immediately became clear to me that this was easier said than done, as I would like to be objective and accurate in the transcription. Should I provide a phonetic transcription? For each sentence over a hundred tokens were examined in the study (several repetitions in different pragmatic contexts by several speakers). Doing a phonetic transcription would thus require making some decisions as to how to solve the variability present in the data. To give just one example, one of the experimental sentences started with the word mollako ‘of the pier’. In this word the k was usually

* For comments, I am grateful to Lourdes Aguilar, Bob Ladd, Pilar Prieto, Daniel Recasens and the audience at WCCFL23. All errors are mine.
voiceless, but sometimes is was voiced, and some other times partially voiced. How should this segment be transcribed? Should I choose the most frequent production perhaps?

It would seem that by choosing a phonemic transcription instead this problem would be solved. But phonemization would necessarily be subjective in some respects. For instance, something that is controversial in Basque is the phonemic status of some palatal consonants. Should I transcribe mollako as /mo₃ako/ or as /moilako/? In the relevant Basque dialects /l/ palatalizes after /i/ and palatal glides are absorbed by a following palatal lateral, so that both phonemic inputs would result in the same output. An alternative orthography is indeedmoilako.

There are two conclusions to be drawn from this Basque example. First of all, as pointed out by Pierrehumbert, Beckman and Ladd (2002) and Ladd (2002), “systematic phonetic representations” have no scientific validity, other than as practical shorthand and allowing for some arbitrariness (see also Beckman 2003). Secondly, the phonemic/phonological level (contrastive units) is also ill-defined in some respects. Here I will concentrate on this second issue.

We may ask now if my difficulties at the phonemic level were due to the fact that there is relatively little work on Basque phonology. I do not think this is the reason. Problems of phonemic analysis arise even in the best studied languages. This is a pervasive phenomenon. The fact is that quite often a phonemic transcription or orthography almost works.

We may note in passing that the problems of arbitrariness or indeterminacy are much greater at the level of underlying representations (URs) in a generative analysis. For instance, is righteous really /rixt+i+i+əs/ (as claimed in SPE)? Do cow and beef share the same UR? (< PIE gwou-). The added problem is of course the requirement that “morphologically related” words should in principle share the same UR for shared morphemes in a generative analysis. Deciding when two words are “related” implies a clear element of subjectivity.

Here I will leave morphophonemics aside and concentrate on the classical phonemic level (which in practice is recognized by most phonologists). The point that I want to make is that even in the case of a fairly well studied language such as Spanish the phonemic status of certain oppositions is less than clear. I will maintain that the possibility of having quasi-phonemic contrasts is inherent to the nature of linguistic categorization. Rather than trying to sweep this issue under the rug or trying to decide on unclear aspects of the phonemic inventory on the basis of the cleverness of competing analyses, I believe we should recognize the emergent character of phonological categorization (Bybee 2001).

Spanish has the phonemic inventory in (1), where phonemes in parentheses are found only in some dialects (but their status in the dialects
where they occur systematically is not disputed). Bolded phonemes, on the
other hand, are of questionable status: some authors include them in the
inventory and some other authors do not.

(1) Spanish phonemes: Questionable phonemes are bolded. Phonemes in
parentheses are found only in some dialects

Consonants: p t tj k b d j g f (θ) s x m n jn l (ʎ) t ɾ

Vowels: i e a o u , Glides: j w

As we can see in (1), the controversial phonemes are the glides, the
voiced palatal obstruent ɾ and the trill ɾ. For different generative analyses
see, among others, Saporta and Contreras (1962), Harris (1969, 1983),
Cressey (1972), D’Introno et al. (1995) and Whitley (1995, 2002).¹ For a
Praguean structuralist perspective see Quilis and Fernandez (1983), Alarcos

Theoretical allegiance does not determine phonemic analysis. For instance,
D’Introno et al. (1995) adopt a generative formalism but agree with the
standard Praguean analysis in recognizing the phonemic status of j and ɾ
but not that of glides, and among generative analyses we find a wide
diversity of opinions. The status of glides and the palatal fricative were the
topic of controversy within American structuralism as well (see Stockwell,
Bowen and Silva-Fuenzalida 1956, Saporta 1956).

In the next three sections I will examine the status of glides, the voiced
palatal obstruent and the trill separately, after which some general
conclusions will be proposed.

2. Phonemic glides?

In Spanish there are some (near-)minimal pairs that appear to rely on a
high vowel vs. glide contrast, as we see in (2):

(2) high V vs. glide contrasts in Spanish (some dialects)

1. Even though the segments included in generative underlying representations
may be more abstract than the phonemes of other theories it is generally possible to
compare across frameworks in the relevant respects because, in practice, generative
analyses also employ phonemes. For instance Harris (1983) provides the phonemic
inventory of Spanish on p. 138, note 6. As Ladd (2002) points out the ‘output of the
lexical phonology’ in some versions of generative phonology corresponds most
closely to the phonemic level of structuralist analyses.

2. Some of these authors include a “fricative” w (with some added diacritic) in
the inventory.
In spite of the existence of these (near-)minimal pairs in at least some varieties of Spanish, many (most?) Spanish phonologists analyze the glides not as independent phonemes but as allophones of the high vowels. This is in part because both configurations in (2) do not have the same weight in the lexicon. Hiatus sequences of the *i.á* type (left column) are a marked or exceptional configuration (and even more so in sequences of falling sonority such as *reiré* ‘I will laugh’). The pattern on the right column of (2), on the other hand, represents the unmarked, regular, situation in the language. Another complication is that all words in the hiatus class also allow a pronunciation with a diphthong under conditions of weak phrasal stress. The contrast is thus one between a regular class of words with obligatory diphthong and an exceptional class of words with possible hiatus (see Aguilar 1999, Hualde and Prieto 2002).

For the most part high vowels and glides are in complementary distribution. That is, for the most part, VV syllabification is predictable in Spanish, as summarized in (3):

(3) Syllabification of vowel sequences in Spanish
   a) If both vowels are [-hi]: hiatus; *te.átró* ‘theater’, *bo.áto* ‘pomp’.
   b) If there is a stressed [+hi] V: hiatus; *marí.a* ‘María’, *ganó.tí.a* ‘hook’.
   c) Otherwise: diphthong; *djénte* ‘tooth’, *dwélo* ‘duel’, *márjó* ‘Mario’, *italjáno* ‘Italian’, *páwsa* ‘pause’. Also across word-boundaries *mjamígo* ‘my friend’ (vs. *miprímo* ‘my cousin’).

From this distribution we can conclude that glides can be considered allophonic variants of high vowels when adjacent to another vowel and not bearing the stress on their own. The explanation for the surface contrast in (2) is that rule (3c) has exceptions (e.g. *kli.énte* ‘client’, *du.éto* ‘duet’).

A general sense for the distribution of exceptional hiatus sequences in Castilian Spanish can be obtained by examining Quilis and Fernández’s (1985: 190-192) transcriptions of “The north wind and the sun”, which these authors offer both in broad phonetic and phonological (Praguean) transcription (a narrow and semi-narrow phonetic transcription are also provided, which do not differ from the broad phonetic transcription in relevant details). In (4) we give all examples of sequences containing a high vowel that are found in that text:

As can be seen, these authors do not include glides in phonological transcriptions. All phonetic glides correspond to phonological high vowels in the context of (3c). There is however one anomaly: the word porfiában which is given as containing a phonetic hiatus, even though it should have a diphthong by rule (3c).³

This text is representative of the general situation in the language: What needs to be lexically marked is that some high vocoids remain syllabic in contexts where they should be realized as glides (Roca 1997, Hualde 1997, Harris and Kaisse 1999, etc). Glides are predictable/allophonic/regular realizations of /i/, /u/ in VV if not stressed (by rule (3c) above). What is “irregular” is the exceptional presence of syllabic /i/, /u/ in the environment of the general gliding rule in some words. That it, although there is is a [i]-[j], [u]-[w] surface contrast, this contrast is not adequately captured by postulating two pairs of phonemes, high vowels and glides, as in (5a). A more adequate characterization of the facts is that in (5b-c):

(5) Analysis of high vowel vs. glide contrast
   a. Problematic phonemization: /i/, /j/, /u/, /w/
   b. Possible phonemization: /i/, /i*, /u/, /u*/
      where /i*, /u*/ differ from /i/, /u/ in not being subject to
distributional rule/constraint (c):
   c. Gliding rule/constraint
      i/u → j/w if adjacent to another V and not stress-bearing.
      e.g. /duélo/ [duélo] vs. /du*éto/ [duéto]

The phonemization in (5b-c) accounts for markedness facts better than the /i/ vs. /j/ analysis in (5a), but it is still not completely satisfactory, as it also fails to capture some facts. First of all, exceptions (with hiatus) are not randomly distributed in the lexicon. Secondly, the diphthong/hiatus contrast has quasi-categorical aspects: it is clear in some contexts, nonexistent in

3. Martínez-Celdrán et al. (2003) offer another transcription of this text where porfiában is transcribed with a diphthong. I believe Quilis and Fernández’s transcription is more representative of careful Castilian Spanish pronunciation in this respect (as mentioned in the text these hiatus sequences can also be reduced to diphthongs in contexts of weak stress).
other contexts and marginal/unclear in yet other contexts. In addition, there
is a fair amount of dialectal and idiolectal variation in this respect (cf. high-
mid vs. low-mid vowel contrast in French and Italian, Tranel 1987, Ladd
2002, etc).

Interestingly, the quasi-phonemic nature of the facts is recognized by
the Spanish Academy in its orthographic rules. Although the diphthong vs.
hiatus contrast is not directly signaled in Spanish orthography, the correct
syllabification of these sequences needs to be determined for the application
of the rules of orthographic accent. One such rule states that stress is
indicated with an accent mark in oxytones ending in a vowel, n or s
(Panamá, revolución, anís). But a systematic exception is made for
monosyllabic words (fe, son, tos). Given this, should one write guión or
guio? It would depend on whether the word has one or two syllables.
According to the most recent rules of the Spanish Academy in cases like
these you may use an accent mark if you strongly feel that the word has a
hiatus and, thus, is bisyllabic (RAE 1999: 46). The existence of inter-
speaker variability and differences in the strength of categorical intuitions
in this respect is thus acknowledged. This is very different from other
orthographic rules. For instance, Spanish speakers are not told to write an
accent mark on sílaba only if they strongly feel that the stress falls on the
antepenultimate. Universal agreement (or clear intuitions) among native
speakers is assumed regarding such phonological matters.

2.1. Distribution of exceptional hiatus in Castilian Spanish

Subject to this variation, exceptional hiatus is far from being randomly
distributed in the lexicon. We may distinguish two classes of words with
morphologically justified hiatus: hiatus due to paradigm effects and hiatus
due to an intervening morpheme boundary (see Navarro Tomás 1977: 158-
159). The hiatus in porfiában ‘they disputed’ finds its justification in other
forms of the same verbal paradigm where the high vowel is stressed and
therefore the sequence must necessarily be syllabified in hiatus, such as
porfián ‘they dispute’ (cf. also: řío ‘river’→ říáda ‘high waters’, dú.o
→ du.éto). In an example such as bokiántfo ‘wide-mouthed’, on the other
hand, the presence of a compound boundary blocks the syllabification of
the sequence as a diphthong (also in prefixation. bíénjo ‘biennium’; and
with some suffixes: xesu.́ita ‘jesuit’).

In addition, there are cases of exceptional hiatus where neither of these
morphological conditions holds, but these tend to occur in some specific

4. “En este caso es admisible el acento gráfico, impuesto por las reglas de
ortografía anteriores a estas, si quien escribe percibe nítidamente el hiato.”
phonological contexts. In particular, hiatus is favored with the sequences ia, io (and to some extent ui) in initial position and where the stress is either on the second vowel of the sequence or on the next syllable. The initiality condition is demonstrated by examples like liána vs. italjána (*italiána); bi.ólogo ‘biologist’ vs. radjólogo ‘radiologist’; di.ána ‘target’ vs. medjána ‘medium, fem.’. Regarding the stress condition, cf.: di.álogo ‘dialog’, dialógo ‘I converse’ vs. djalogó ‘s/he conversed’; dú.o, du.ál, du.alísmo ‘dualism’ vs. dvalidád ‘duality’. There is no exceptional hiatus after the stress: kópja ‘copy’, istója ‘history’ (*kópi.a, *istóri.a) (see Hualde 1997, 1999).

Given this biased distribution it should be clear that the analysis in (5b-c) does not adequately account for all the facts.

2.2. Historical origin of diphthong/hiatus contrast

There are two main possible origins for sequences of rising sonority in Spanish:

a) On the one hand, the diphthongs jé, wé arose from the breaking of stressed low-mid vowels: terra > tjéra ‘land’, porta > pwértta ‘door’. Sequences with this origin are unexceptionally pronounced as diphthongs. Since most of the words with ie, ue have this origin, this fact explains why hiatus is rare with these particular sequences.

b) Secondly, in Latin we find heterosyllabic vowel sequences with unstressed high vowels. These sequences were also created by the deletion of certain intervocalic consonants. The general tendency in the language has been the reduction of the hiatus to a diphthong (Quilis 1993, Lloyd 1987, Penny 2002): pretiu(m) > prétjo ‘price’, Italia > itálja ‘Italy’, rugitu(m) > ru-ido > rwido ‘noise’. But contraction has exceptionally been blocked and the hiatus has been preserved in some of these words: diabolu > di.áblo ‘devil’, cliente > kli.énte, crudelitate > kru.eldád (hiatus possible in these words for some speakers).

From a diachronic point of view, the exceptions to the gliding rule in (5c) are words where the historical tendency i.a > ja has been blocked. Hiatus has been variably preserved: (a) when morphologically supported and (b) in “strong” positions = initial position and not too far before the stress (These hiatus words may then have acted as analogical attractors for borrowings and neologisms meeting the phonological conditions).

What needs to be explained is thus why contraction has been blocked in specific contexts. For morphologically-justified exceptions, the

5. For Catalan, see Recasens (1993), Cabré and Prieto (to appear).
explanation is more or less clear (on paradigm effects, cf. Steriade 2000). But what about cases without morphological justification? To repeat, almost always these exceptional hiatus words contain sequences of rising sonority which meet the two conditions of initiality and proximity to the stress (before it). These are necessary but not sufficient conditions; rather, contrast is possible under these conditions. Why is the preservation of historical hiatus favored in these contexts?

2.3. Explaining the distribution of exceptional hiatus

Hualde and Chitoran (2003) tested the hypothesis that the diachronic tendency to reduce original rising hiatus sequences to diphthong has been exceptionally blocked in positions where vowels independently have relatively greater duration, because of the prosodic or rhythmic patterns of the language.

In a first experiment, 4 Spanish speakers from Spain (Sp4 = author JH) read a randomized list of words containing a ia sequence, within a carrier phrase (4 repetitions). For all test tokens the duration of the sequence ia was measured (using PRAAT). To test the effect of stress on the duration of the sequence ia was measured (using PRAAT). To test the effect of stress on the duration of the sequence, words starting with the sequence di- were classified into three groups, depending on the position of this sequence with respect to the stress: stressed (diáspora, diáceno, diána), pretonic (diamante, diafragma, diatriba, diabétes, diagráma) and prepretonic (diapasón, diagonál, diamétral).⁶ Means in milliseconds for each position and speaker are shown in Table 1.

<table>
<thead>
<tr>
<th></th>
<th>Sp1</th>
<th>Sp2</th>
<th>Sp3</th>
<th>Sp4</th>
</tr>
</thead>
<tbody>
<tr>
<td>Stressed</td>
<td>192.61</td>
<td>161.74</td>
<td>136.77</td>
<td>165.87</td>
</tr>
<tr>
<td>Pretonic</td>
<td>154.19</td>
<td>94.574</td>
<td>108.19</td>
<td>127.85</td>
</tr>
<tr>
<td>Prepretonic</td>
<td>118.33</td>
<td>95.10</td>
<td>99.38</td>
<td>107.09</td>
</tr>
</tbody>
</table>

An ANOVA and post hoc comparisons revealed that stressed sequences are significantly longer than the others for all 4 speakers. In addition, for 2 of the 4 speakers, Sp1 and Sp4, pretonic sequences have

---

⁶ The complete test list contained also words with other vowel sequences. A parallel experiment on Romanian produced similar results, see Chitoran and Hualde (2002), Hualde and Chitoran (2003).
significantly greater duration than those in words where the stress is further to the right. (For the other two speakers pretonic and prepretonic sequences do not significantly differ in duration.)

This experiment does not allow to tease apart lexical and purely phonetic effects, since some of the words may have contained lexical hiatus for some of the speakers. For this purpose, we need to examine monophthongs in the same positions. In a second experiment, Hualde and Chitoran (2003) examined the duration of the first vowel in five triplets differing in the position of the stress, e.g. número/número/numeró, in a carrier phrase (3 speakers, 4 repetitions). The results indicated that whereas the main stress effect on duration is that stressed initial syllables (in proparoxytones) are longer than unstressed initial syllables, there is also a tendency for initial syllables immediately before the stressed syllable (in paroxytones) to be longer than initial syllables further away from the stress (in oxytones). The observed duration cline is thus the following: stressed \( \rightarrow \) pretonic \( \rightarrow \) prepretonic. This is also consistent with results for Catalan reported in Recasens (1991b).

We may thus conclude that the perception of hiatus is possible only in positions with enough durational substance. In these positions historical recategorization of \( \text{i.a} \) as a diphthong has been blocked in some lexical items. The same basic durational effects are found with monophthongs (\( \text{número} > \text{númeró} > \text{numeró} \)), but they remain below the level of awareness because there is no possible contrast.

Simonet (2003) speculates that intuitions about syllabification in hiatus should be more robust or consistent across speakers in stressed than in (immediately) pretonic position, given the durational difference between tonic and pretonic syllables. He conducted a paper and pencil test where participants were asked to syllabify 45 test words (including both nonce and real words) indicating how confident they were about the chosen syllabification in a 5-point scale where 1 represented clear hiatus and 5 clear diphthong. For instance, for the nonce word \text{miabóco} \text{ subjects were given the options mi.a.bo.có = na.na.ná and mia.bo.có = na.na.ná. All test items had initial ia type sequences and differed in the position of the stress: Type A = sequence stressed (\text{miáboco}, \text{diálogo}), Type B = sequence pretonic (\text{miábóco}, \text{diálogó}) and Type C = sequence prepretonic (\text{miabocó}, \text{dialogó}). The results (from 12 speakers of Peninsular Spanish) were consistent with the hypothesis. The number of diphthong responses increases with distance from the stress. Averages were: Type A = 2.1 pts., Type B = 3.4 pts. and Type C = 4.3 pts. Thus, participants showed a clear preference for hiatus syllabification in words of the \text{miáboco} type, for diphthongs in the \text{miabocó} type and had less clear intuitions regarding the \text{miabóco} type. Simonet’s interpretation is that the difference between diphthong and hiatus is not categorical, but gradient, and depends on
general gradient patterns of prosodic lengthening. There is no precise
duration point where a vowel sequence would be considered a diphthong
instead of a hiatus. What we have is a gradient difference that goes from
prototypical long hiatus, to prototypical short diphthongs. There are two
categories, hiatus and diphthong (or, equivalently, high vowel and glide),
but no precise boundary between the two.

The two recent studies reviewed in this section provide a partial
explanation for why certain positions are more likely to preserve a lexical
hiatus (stress condition): the longer duration of segments in those positions
(due to independent rhythmic factors) favor the categorization as hiatus of
sequences of vocoids. The reason for the word-initiaility condition (the fact
that word-internal exceptional hiatus is possible only if paradigmatically
supported) remains to be investigated.

2.4. Conclusion regarding glides

Paradoxically, although all glides are contextual allophones of the high
vowels in Spanish, there is a contrast between high vowel and glide in
certain positions.

There are surface contrasts because in a relatively small number of
cases and in very limited contexts there is exceptional syllabification as
hiatus of sequences that according to the general pattern of the language
should be expected to be realized as diphthongs (the historical tendency to
reduce hiatus to diphthong has been blocked). Exceptional hiatus sequences
are not randomly distributed in the lexicon and, in fact, in some specific
positions “exceptional” hiatus is the preferred configuration (in some
varieties of the language).

3. The voiced palatal obstruent /j/

The voiced palatal obstruent j is another segment of disputed phonemic
status. This sound corresponds to the underlined letters in examples such as
paranoja, cebolla ‘onion’ (in the majority yeísta pronunciation).7 In
Castilian Spanish this segment is realized with a broad range of constriction
degree. In an acoustic study, Aguilar (1997: 69-73) distinguishes four
allophones, all of which are said to occur in free variation in word-initial

7. Yeísmo refers to the delateralization of the palatal lateral phoneme /ʎ/,
    orthographically represented as ll.
position: approximant [j] ~ fricative [ʃ] ~ affricate [dʒ] ~ stop [t]. Arguably this allophonic categorization in 4 types is imposed by the IPA. Perhaps it would be more accurate to speak of a continuum of realizations regarding constriction degree. Speakers of this and many other Spanish dialects have trouble with English minimal pairs such as *Yale* and *jail*, since they do not perceive the initial segments in these words as categorically different.

The only context where [j] is found is in syllable-initial position. Glides cannot be strengthened if preceded by a consonant in the same syllable. Thus, for instance, *italiano* ‘Italian, masc.’, always with a pure glide, may contrast with *italiano* (y tal llano) ‘and such a plain’ where the element transcribed as [ʃ] has the range of realizations described above.

In principle it would appear that [ʃ] can be analyzed as resulting for the (variable) reinforcement of [j] in syllable-initial position, and, therefore as an allophonic variant of *[ʃ]*:

(6) Syllable-initial glide fortition

\[
{ʃ} → (ʃ) (¬[ʃ] ~ [dʒ] ~ [j]) / V
\]

A well-known problem for this analysis is the fact that there are some surface contrasts between [ʃ] and [ʃ]:

(7) [ʃ] vs. [ʃ] contrasts

- **desjérto** ‘desert’ vs. **des.jélo** ‘thawing’
- **abjérto** ‘open’ vs. **ab.jékto** ‘abject’,
- **bonjáto** ‘yam’ vs. **kón.juxe** ‘spouse’, **kón.jéba** ‘it implies’

For some phonologists, these facts constitute enough evidence to conclude that [ʃ] is an independent phoneme in Spanish. For other linguists, however, these contrasts are not completely decisive, since they depend on syllabification, which is for the most part predictable from morphological structure: in the examples with glide fortition there is a prefix **des-**, **ab-**, **kón-**, which introduces a boundary for syllabification (although this morphological analysis is not so clear in **kón.juxe**, since there is no

---

8. Martínez-Celdrán and Fernández-Planas (2001) argue that a more appropriate symbol for the affricate would be [ʃ], since its place of articulation is palatal.
independent morpheme \( \text{ju}x \) and, furthermore, under this morphological analysis this word would have the stress on the prefix, contrary to the normal situation). 9

There are in fact two main positions regarding the status of \( j \) in Spanish: (A) It is a consonantal allophone of the vowel /i/ (in syllable initial prevocalic position) and (B) It is not the same phoneme as the vowel /i/, because it is a consonant.

Spanish dialects vary widely in the range of realizations of the sound we have been transcribing as \( j \). One extreme is represented by Argentinean Spanish. In this variety \( j \) has undergone further fortition: \([j]\) > \([z]\) > \([f]\). Rather than being optional, as in Castilian, in Argentinean consonantal realizations are obligatory. For instance, whereas in Castilian yo ‘I’ may be realized as \([j]\) ~ \([j]\) ~ \([j]\) depending, perhaps, on degree of emphasis, in Argentinean only obstruent allophones are found: \([z]\) ~ \([s]\). On the other hand, in Argentinean Spanish the production in this consonant in words with orthographic hiV is stigmatized as indicating ‘low educational level’. In Standard Argentinean Spanish now \([z]\) ~ \([s]\) contrasts with \([j]\). We find \([z]\) ~ \([j]\) in yeso, llena, tramoya, cebolla but \([j]\) ~ \([j]\) in words spelled with (hi) such as hiel, hiena, paranoia. There is even a minimal pair from what originally was a single word with two alternative spellings: yerba \( \text{zérba} \) ‘mate’ vs. hierba \( \text{jérba} \) ‘grass’. The phonemic link between \(/\text{j}y/\) \([z]\) ~ \([s]\) and \(/\text{i}/\) \([i]\) ~ \([j]\) has been broken (pace Harris and Kaise 1999). 10

The interesting thing for our purposes here is that other dialects without obligatory fortition also have an orthographically-based (but nonetheless real) ‘quasi-contrast’. The reluctance of educated speakers to strengthen the pronunciation of words with orthographic (hi) is by no means an Argentinean idiosyncrasy. In many other areas as well, words spelled with hie- are rarely pronounced with a strong fricative or noncontinuant consonant by educated speakers (see, for instance, Navarro Tomás 1977: 50, Dalbor 1997:217). The greater the phonetic distance between the strongest realization of \( j \) and a pure palatal glide in the dialect, the greater the likelihood of speakers establishing a separate category for words spelled with a vowel.

---

9. The explanation for this word is found in its etymology. Although its structure is opaque in Spanish, \( \text{cónyuge} \) does in fact derive from a Latin prefixed form.

10. At the morphophonemic level we find alternations like in \( \text{léj} \) ‘law’, \( \text{légés} \) ‘laws’ in Argentinean Spanish.
As noted above, for Castilian Spanish Aguilar (1997) finds realizations of word-initial \( \text{j} \) that range from a glide to a plosive. But this is compatible with the fact that different words allow a different range of realizations along the constriction parameter.

In my own lexicon I believe I have 3 word classes along the relevant parameter. Assignment of words to a class or another is in part orthographically driven and in part dependent on word frequency:

(8) Word classes and range of realizations (author’s Castilian variety)

\[
\begin{align*}
\text{hiato} \ '\text{hiatus}' & \quad \text{i.á} \sim \text{já} \\
\text{hiena} \ '\text{hyena}' & \quad \text{jé} \sim \text{jé} \\
\text{yema} \ '\text{yolk}' & \quad \text{jé} \sim \text{jé} \sim \text{jé}
\end{align*}
\]

This three-way distinction with overlap of phonetic ranges is the result of the convergence of two quasi-contrasts: the high-vowel vs. glide contrast (i.iá vs. já) and the glide vs. palatal obstruent contrast (já vs. já). This would seem to be a case of “word-specific phonetics” (Pierrehumbert 2002) where awareness is greater than usual because of the quasi-categorical status of the contrasts involved.

4. Tap \( \text{r} \) vs. trill \( \text{ř} \): one phoneme or two?

Spanish has a robust contrast between rhotic tap and trill, for which minimal pairs can easily be provided: \( \text{káro} \ '\text{caro}' \ 'expensive' vs. \( \text{kářo} \ '\text{carro}' 'cart, car'; \( \text{pěra} \ '\text{pear}' vs. \( \text{pěřa} \ '\text{dog, fem.'}; \( \text{kería} \ '\text{I wanted, imp.'} \) vs. \( \text{kěřía} \ '\text{I would want’}. The reason why the phonemic nature of tap and trill as independent phonemes has been questioned is that the contrast is limited to the word-internal intervocalic position. Elsewhere the contrast is neutralized: in some positions only \( \text{ř} \) is found; in other positions only \( \text{r} \); and in yet other positions there is an indistinct/variable \( \text{r} \) sound.\(^{11}\)

(9) Distribution of rhotics in Spanish

\[\text{a) Only } \text{ř} :\]

- Word-initial
  
  \#___ řóka ‘rock’, a řóma ‘to Rome’

- After a heterosyllabic consonant

---

C. al. ředědóř ‘around’, en. ředo ‘mess’, is. řaelita ‘Israeli’

b) Only r:
- After a tautosyllabic consonant
  C__ bróma ‘joke’, grámo ‘gram’
- Word-final before a vowel
  V__#V sér amígos ‘to be friends’
- Indistinct/variable r:
  - Before a consonant
    V__ C párte ‘part’; V__#C ser poéta ‘to be a poet’
  - Before pause
    V__## (sér o nó) sér ‘(to be or not) to be’

An additional morphophonological restriction is that word-final r always corresponds to r, never to ř, before a vowel in morphologically related words. Thus, there are sets of related words such as olór ‘smell’, olóres ‘smells’, oloróso ‘smelly’; sepór ‘gentleman’, sepóres ‘gentlemen’; seþóra ‘lady’; but not olór, *olór es, *oloróso, *seþór, *seþóres, *seþóra

(With very few exceptions, which, furthermore, have never been noticed before in the literature on this topic, see below).

Two alternative analyses have been proposed:


Although Harris (2001) refers to the one-phoneme analysis as “the generative consensus” and “the standard analysis”, from the references above it should be clear that there is really no consensus on this analytical matter. Authors proposing two phonemes need to account for the limited nature of the contrast, regarding its phonological environment. Under the one-phoneme analysis the problem is obviously how to account for the tap/trill contrast in intervocalic position.

Harris’ one-phoneme proposal essentially is that the contrast between, e.g., péro (pero) ‘but’ vs. pěro (perro) ‘dog’ is underlingly a contrast between single and geminate segment. Harris (2001) proposes the underlying representations /peRo/ and /peRRo/, where R is an underlying segment “unspecified for the features that distinguish [r] from [ř] [= our ř] in phonetic representations” (Harris 2001:136). From these representations surface forms are obtained by the application of a series of ordered rules.
One possible objection to this analysis is that Spanish does not have morpheme-internal geminates (there are only a couple of exceptions with /bb/ and /nn/: obvio, perenne). I suppose an answer to this objection would be to state that the true generalization is instead that Spanish does not have geminates except for (underlying) geminate rhotics.

A crucial element of the analysis, since it is the basis for a step in the derivation of the trill, is the claim (already present in Harris 1983) that tap+trill sequences across word boundaries are non-distinct from single trills; that is, salí rápido ‘I left quickly’ = salir rápido ‘to leave quickly’; “[r] alone is not distinct from [ɾ-], unless obviously, the two segments are deliberately separated by an artificial pause or naturally fall into separate phonological phrases. I believe this is true for all dialects, whatever the range of possible realizations of [r] and [ɾ] individually may be.” (Harris 1983: 63).

Harris is not alone in holding this view, which is also expressed in Quilis and Fernández (1985:148). Pensado (1999), on the other hand, appears to assume that the tap+trill sequence is different from a single trill, but the point is not made very clearly.

Harris’ and Quilis and Fernández’s hypothesis is a sensible one, since a word-internal trill may be realized with two, three or more occlusions without this creating any phonological contrast. That is, caro, with a tap, contrasts with carro, with a trill, but carrrro, with a longer trill, can only be an emphatic rendition of carro, not a different word. Nevertheless, regarding inter-word sequences, my intuition is different from Harris’ and Quilis and Fernández’s. In my opinion in examples such as salí rápido vs. salir rápido the contrast may be neutralized, but it may also be preserved in natural speech. My hypothesis is that across word-boundaries there is indeed a three-way contrast: tap vs. trill vs. r+trill. I believe this also holds for sequences created in prefixation:

(11) 3-way rhotic contrast (hypothesized)
   a. Across word boundaries:
      r dar ocas ‘to give geese’ vs.
      ṭ da rocas ‘s/he gives rocks’
      rr dar rocasi to give rocks’
   b. Across prefix boundaries:

Furthermore, the number of words with a geminate does not increase significantly if we count all word-internal geminates, not only morpheme-internal ones, since additional examples are only those created by transparent prefixes like in- (innato) and sub- (subversión).
To decide on this difference of opinion I conducted the experiment described in the following subsection.

4.1. Production experiment: Vr+trill vs. V+trill across word boundaries

4.1.1. Materials, subjects and analysis

20 contrasting sentences were constructed to test the production of word-initial r and the inter-word sequence r#r. Care was taken to construct the sentences in such a way that a prosodic break would not be expected between the two relevant words (e.g. Aunque salí rápido, no llegué a tiempo ‘Even though I left quickly, I did not arrive in time’). The 20 experimental sentences form 5 sets of 4 sentences, containing 2 pairs differing only in the order of main and subordinate clause.

Together with these sentences, other examples contrasting n and n#n as well as s and s#s were also constructed and read by the participants in the experiment to provide some comparison regarding the duration of one vs. two identical consonants in a row. The complete list of examples is given in the appendix (They were presented in this order with a blank line every 4 sentences).

Subjects were asked to read the list twice at a comfortable rate. This was recorded directly onto a personal computer. 3 female native speakers of Spanish from Spain participated in this experiment.

The recorded data were analyzed with PRAAT. The measurement that was taken was the duration of the target segments. Counting occlusions for rhotics was determined not to be practical because trills may be realized with incomplete occlusions (as fricatives or approximants, Lindau 1985, Blecua 2001, Hammond 1999).

4.1.2. Results and discussion

Average durations in ms. (with standard deviations in parentheses) for each speaker are given in Table 2. The table also includes the durations that were obtained for single and double n and s. Rows: r1 = r, r2 = r#r, n1 = n.

13. These two examples also differ in their basic stress pattern, but this difference can be eliminated either by placing stress on the initial member of the compound (súper-rápido) or by deaccenting the verb in supe rápido.
n2 = n#n, s1 = s, s2 = s#s. The difference in duration between r1 and r2 is approximately of the same magnitude as that between s1 and s2, but not as great as that between single and double n. The durational contrast between r1 and r2 is greatest for Spkr1 (over 30 ms in the means) and smallest for Spkr3 (about 16 ms). Nevertheless, for all 3 speakers the duration of r#r sequences was significantly greater than that of single r; 2-tailed t-test, all p <.0001.

<table>
<thead>
<tr>
<th></th>
<th>N</th>
<th>Spkr1</th>
<th>Spkr2</th>
<th>Spkr3</th>
</tr>
</thead>
<tbody>
<tr>
<td>r1</td>
<td>20</td>
<td>59.13 (12.3)</td>
<td>61.11 (15.1)</td>
<td>49.55 (6.8)</td>
</tr>
<tr>
<td>r2</td>
<td>20</td>
<td>90.35 (21.0)</td>
<td>83.70 (17.5)</td>
<td>65.43 (12.7)</td>
</tr>
<tr>
<td>n1</td>
<td>8</td>
<td>72.35 (13.5)</td>
<td>75.01 (7.1)</td>
<td>74.36 (8.3)</td>
</tr>
<tr>
<td>n2</td>
<td>8</td>
<td>127.05 (17.4)</td>
<td>126.03 (7.6)</td>
<td>124.19 (15.5)</td>
</tr>
<tr>
<td>s1</td>
<td>8</td>
<td>122.32 (18.8)</td>
<td>90.98 (15.0)</td>
<td>94.28 (6.8)</td>
</tr>
<tr>
<td>s2</td>
<td>8</td>
<td>136.17 (16.3)</td>
<td>111.72 (22.0)</td>
<td>123.86 (13.3)</td>
</tr>
</tbody>
</table>

To the extent that the one-phoneme analysis depends on salí rápido = salir rápido, the data from this experiment do not offer support for this analysis. Rather, here we have a case of incomplete neutralization. In normal speech the two sequences can be identical but, clearly, they can also be different. In addition to the difference in duration of the consonant reported here, there are likely to be differences in the duration and formant structure of the preceding vowel (with a shorter vowel and more coarticulation or “r-coloring” in a closed syllable). We leave this for future research.\(^{14}\)

We are thus left with the two-phoneme analysis: r vs. ř. What needs to be explained is why these two phonemes contrast only in such a limited phonological environment.

4.2. The history of R

The limited contrast between tap and trill has an obvious historical explanation: it derives from an earlier contrast between single and geminate

\(^{14}\) To definitely establish the facts it would be useful to conduct a perception experiment as well. This was not done here because it seemed clear to me that some of the production tokens were more ambiguous than others. I concluded that depending on which token was selected the outcome of the perception test was likely to be different. To be able to obtain replicable results, this issue must be dealt with in some objective manner.
consonant. That is, the one-phoneme analysis is correct for an earlier stage.
As is well known, in Latin a single vs. geminate contrast existed for almost all consonants. The contrast was found only in word-internal intervocalic position. In the evolution from Latin to Spanish, geminates became single consonants, -CC- > -C-, e.g., bucca > boca, gutta > gota, except that -II- > ∊, -nn- > ñ and -rr- > ð (Menéndez Pidal 1973, Lloyd 1987, Penny 2002).
The Spanish trill is not a geminate since it syllabifies as onset with the next vowel (on the articulatory differences between trills and taps see Recasens 1991a, Recasens and Pallarès 1999). In Italian, where the single/geminate contrast of Latin has been preserved, carro is syllabified as car-ro (see, e.g., Muljačić 1972), whereas in Spanish the syllabification is ca-ro (this difference between the two languages should in principle have phonetic effects, including in the duration of the preceding vowel, see Maddieson 1985).

The limited context for the contrast between tap and trill in Spanish is thus explained by its origin in a single vs. geminate opposition. There has been no regular sound change creating new instances of intervocalic trills. The distribution between tap and trill in Spanish is, nevertheless, not exactly the same as that between single and geminate r in Latin because of independent evolutions in other positions. In particular, in Spanish and some other Romance languages there has been fortition of word-initial r-; e.g. Lat. romána > Sp. romána (in other Ibero-Romance languages I- and n- also underwent fortition).

Since the intervocalic trill derives from a Latin geminate, it is thus not surprising that there are no proparoxytonic words of Latin origin in Spanish with a trill in the onset of the final syllable, which, as a geminate, would have made the penultimate heavy in Latin, attracting stress. The only words with this pattern are of onomatopoetic (cháncharras-máncharras ‘pretexts’), Basque (Chávarri, Achúcarro) or unknown origin (tábarro ‘type of wasp’).

The neutralization of word-final rhotics (e.g. olor/olores but not *olor/olorres) also has an obvious historical origin: If r was already word-final in Latin, it couldn’t be a geminate. Word-final consonants also arose from the deletion of e after a single coronal consonant, as in panepanes > pánepanes ‘bread sg/pl’, amorelamoseres > amorlamoseres ‘love sg/pl’. Since, on the other hand, the few words ending in -rr, such as torre (< Lat.

15. In the future and conditional of querer, vowel deletion has produced a trill from two taps: quer(e)r e > querré ‘I will want’. There are no other examples of this change.
turrem) did not lose their final vowel, there is no *tor/torres (cf. torre/torres).\footnote{Penny (2002: 83), assumes that the final vowel was lost in this case as well and later restored, so that an alternation “probably once existed between singular tor and plural torres (< TURRES)”.
}

Given this evolution, there cannot be any native words of the *olor/olorres type. Such words could only be incorporated to the Spanish lexicon by borrowing from a language with the relevant pattern. Basque is such a language. In Basque most words ending in a rhotic have a trill before a vowel-initial suffix, unlike in Spanish, but there are also a few words with a tap in this context:

(12) Basque word-final rhotics

<table>
<thead>
<tr>
<th>Basque Word</th>
<th>Spanish Word</th>
</tr>
</thead>
<tbody>
<tr>
<td>lur</td>
<td>‘land’</td>
</tr>
<tr>
<td>luña</td>
<td>‘the land’</td>
</tr>
<tr>
<td>luñak</td>
<td>‘(the) lands’ (general case)</td>
</tr>
<tr>
<td>ur</td>
<td>‘water’</td>
</tr>
<tr>
<td>ura</td>
<td>‘the water’</td>
</tr>
<tr>
<td>urak</td>
<td>‘(the) waters’ (exceptional)</td>
</tr>
</tbody>
</table>

In the Spanish spoken in the Roncal Valley of Navarra, which was formerly in contact with Basque, adjectives referring to towns in the Valley take the borrowed Basque suffix -(t)ar, whose final rhotic is a trill before a vowel in the (Spanish) plural or feminine:\footnote{Incidentally, apparently this is not the only gap regarding the distribution of rhotics in Spanish that Basque borrowings have helped to fill. Harris (2001) remarks that the trill is not found after a glide in Spanish words. But this configuration is found in Aurrerá ‘name of a chain of supermarkets’ (from Basque aurrera ‘onwards’), Abaurrea ‘name of a town in Navarra’, aurresku ‘name of a dance’ and the adjective aurragado ‘said of badly tilled land’, of Basque origin, according to the dictionary of the Spanish Academy.
}

\[ burguar/burguiarres ‘person from Burgui sg/pl’, gardiar/gardiarres ‘person from Garde sg/pl’, gardiarra ‘woman from Garde’, etc. \]

There are, thus, monolingual Spanish dialects where a word-final rhotic may surface as either a flap or as trill before a vowel-initial suffix, in a lexically-specified manner.

In conclusion, the tap \( \acute{r} \) and the trill \( \tilde{r} \) are two distinct phonemes in Spanish, whose limited contrast is directly explainable from their origin in a single/geminate opposition. But they are clearly more “closely related” than
other pairs of phonemes, since the contrast is found only in intervocalic word-internal position and is neutralized in an indistinct r sound in preconsonantal and prepausal position.

5. A general view on quasi-phonemic contrasts

Ladd (2002) makes the point that just as categorization in other domains can be at different levels (e.g. among mammals dogs are different from cats and the category dog includes, among others, collie and spaniel), phonological categorization may also involve more and less inclusive categories. He proposes the following schema for mid vowels in French and Italian, to account for the fact that there is a special link between, for instance, /e/ and /e/, which is not found between, say, /e/ and /u/ (as Trubetzkoy already remarked).

(13) Ladd (2002): French and Italian mid vowels

\[
\begin{align*}
\text{E} & \quad \text{O} \\
/e/ & \quad /e/ \\
\end{align*}
\]

The relationship between tap r and trill r in Spanish can be conceived in this manner as well: Spanish has a rhotic category which includes the two categories of tap and trill.

To this view of categorization, I would add that category boundaries may be fuzzy (Taylor 1995). In Ladd’s example, some collies may be different from spaniels, but there are mutts. Mammal vs. non-mammal is usually easier, but there are platypuses. In my own categorization of types of mammals, sheep and goats are different animals, but I have often had the experience of not knowing to which of these two categories of mammals a given animal belonged.

Phonological categories “tend” to be discrete (p or not-p). That’s why phonemic analysis works most of the time. But there are areas of fuzziness probably in every language. If we consider the first two cases discussed in this paper, the ranges of ia, ja and ja show greater or lesser overlap depending on the dialect, the style and the speaker. The extent of the overlap may determine their categorization for a given speaker.
F. de Saussure remarked that “En matière de langue on s’est toujours contenté d’opérer sur des unités mal définies”. That’s the way it should be, because that is the nature of linguistic categories. Language is probabilistic (Bod et al. 2003) and linguistic categories are emerging entities (Bybee 2001).

Appendix: Set of sentences for experiment on rhotics:

La ardilla come nueces y el león cebras.  
No te darán nada aunque lo hagas todo.  
Cuando eres joven no sabes siempre lo que es mejor para ti.  
Aunque tienes sólo diez minutos debes hacerlo despacio.  
Suele dar resultado este experimento.  
Salvo que una vez partí rocas con ella, no he utilizado nunca esa maza de hierro.  
Aunque salí rápido, no llegué a tiempo.  
Cuando era niño yo nunca pedí regalos a mis abuelos.  
Yo no suelo dormir roncando, que yo sepa.  
Salvo para partir rocas con ella, nunca he utilizado esa maza de hierro.  
No llegué a tiempo, aunque salí rápido  
Las ardillas comen nueces y los leones cebras.  
Aunque hagas todo no te darán nada  
Cuando se es joven, no se sabe siempre lo que es mejor para uno.  
Debe hacerse despacio aunque tengas sólo diez minutos  
Este experimento suele dar resultado.  
Yo nunca pedí regalos a mis abuelos cuando era niño.  
Según me dicen, dormí roncando.  
Nunca he utilizado esa maza de hierro, salvo para partir rocas con ella.  
De niño no me gustaba pedir regalos a mis abuelos.  
A pesar de salir rápido, no llegué a tiempo.  
Los leones comen cebras y las ardillas comen nueces.  
Nadie te dará nada aunque hagas todo  
No sabes siempre lo que es mejor para ti, cuando eres joven.  
Aunque uno tiene sólo diez minutos debe hacerlo despacio  
Generalmente da resultado este experimento  
Dormí roncando, según me dicen.  
Debe hacerse despacio aunque uno tenga sólo diez minutos.  
No llegué a tiempo, a pesar de salir rápido  
El león come cebras y la ardilla come nueces.  
Aunque hagas todo nadie te dará nada  
Uno no sabe siempre lo que es mejor, cuando se es joven.  
Este experimento generalmente da resultado.  
No me gustaba pedir regalos a mis abuelos cuando era niño.  
Que yo sepa, no suelo dormir roncando.  
Nunca he utilizado esa maza de hierro, salvo que una vez partí rocas con ella.
References


Padgett, Jaye. 2003. “Systemic contrast and Catalan rhotics”. Univ. of California, Santa Cruz, available at ROA.


