

# Effects of prosodic and segmental variables on vowel sequences pronunciation in Spanish

Lourdes Aguilar

Departament de Filologia Espanyola, Universitat Autònoma de Barcelona

E-mail: Lourdes.Aguilar@uab.es

## ABSTRACT

This paper reports an acoustic study in which Spanish speaker's productions of the vowel /a/ in contact with /a e i/ in word boundaries were obtained during the reading of texts. All tokens were measured for vowel quality (F1/F2) and duration. The obtained data allowed us to classify the observed phonetic results according to phonological processes (hiatus, reduction, monophthongization) and the duration values were statistically processed in order to investigate the effects of some segmental and prosodic conditions on vowel sequences pronunciation. The primary purpose of the study was to provide data for Peninsular Spanish showing the differences between the phonetic outputs of several vowel combinations across word boundaries in an experimental setup.

## 1. INTRODUCTION

It is customary to describe Spanish as a language that do not support adjacency of heterosyllabic vowels ([1], [2]) and the effort is spent on investigating the concatenation-favouring contexts and the occurrence restrictions (see e.g. [3]). It is widely assumed that, where identical adjacent elements occur across word boundaries, one is eliminated in fast speech. Similarly, the adjacency of two non-identical vowels generally surfaces as a monosyllabic group [1]. More precisely, four types of processes that can occur when two or more vowels are in contact between words are describe in [4]: hiatus (-V.V-), reduction (-VV-), diphthongisation (GV, VG), deletion (V) Two identical vowels are always pronounced as a single vowel if the speech style requirements are met: that is, fast but not emphatic speech style. On the other hand, the adjacency of two non-identical vowels can result in a process of reduction, diphthongisation or deletion. The difference between reduction and diphthongisation is only due to the absence or presence of a high vowel, respectively. As far as deletion is concerned, [4] restricted this process to some lexicalised collocations in non-standard sociolects.

With reference to Northern New Mexican Spanish, [5] also based his predictions about the resolution of vowel sequences across word boundaries in the analysis of samples proceeding from casual speech. Nevertheless, it must be noted that besides the speech style, the phenomena

related to adjacent vowels across word boundaries are conditioned by other factors, such as the phonetic environment, the grammatical category and the frequency of occurrence of words, or the given-new information effect. This amount of factors makes a precise phonetic prediction difficult and making generalisations in uncontrolled speech samples can lead to inaccurate results.

From the above considerations, one may wonder to what extent the assumption of anti-hiatus tendency is motivated by the speech style and the large amount of variation in the observed speech samples, at least, at a post-lexical level. In other words, nothing in the literature appears to contradict the hypothesis that recategorization is optional when two content words are found in the speech chain. Some evidences have been reported for other languages ([6]). The purpose of the present study is to provide data for Peninsular Spanish showing the differences between the phonetic outputs of several vowel combinations across word boundaries, and to investigate which variables are affecting the duration variability in an experimental setup.

## 2. EXPERIMENTAL DESIGN

In order to test our hypothesis, a corpus of sequences of two content words was constructed, which has to fulfill certain conditions: the bilabial/dental unvoiced stop consonants were chosen as pre- and post-vocalic contexts, and the syllables in the word were open, thus ensuring that all items appeared systematically in the same phonetic environments. Each word combination had the structure Noun-Adjective: a noun ending with the vowel /a/ is combined with an Adjective beginning with one of the vowels /a e i/. The groups /aa ae ai/ were observed in all the possible lexical stress patterns, that is: [aa áa áá] [ae áe aé] [ai ái aí]. The Noun-Adjective sequences were inserted in texts, in three different prosodic environments: initial, medial and final Intonational Phrase position. In total, the corpus was composed of 36 word combinations.

The final corpus to be read consisted of five texts, to which five filler texts were added. Five male native Peninsular Spanish speakers were asked to read the set of texts, following the Latin square procedure to avoid the presentation-order effect. Any instruction about the speaking rate or the intonation was not given to them. All tokens were processed with the speech analysis software

Praat. Waveforms displays and broad-band spectrograms were plotted for each sequence, and measurements for vowel quality and duration were taken.

### 3. RESULTS

Results are organised according to the two objectives of the study. (a) The observation of the phonetic result of two adjacent vowels across word boundaries. (b) The investigation of the factors affecting the duration variability of vowel sequences between words.

#### 3.1 PHONOLOGICAL PROCESSES

The measurements on F1/F2 and duration allowed us to classify the phonetic results into three categories: hiatus, reduction and monophthongization, defined as follows.

a) Hiatus. The acoustic output is almost identical to the sum of two vowels in an isolated context: i.e. two distinct vowels were clearly observed in the waveform display, even in the case of adjacent identical vowels.

b) Reduction. Below this label, both the reduction of two identical vowels and non-identical vowels are included. In the case of identical vowels, the reduction is manifested by the presence of a long vowel that replaces the segments initially forming the group. The lengthening was established with reference to the mean values of a set of vowels pronounced by the same speaker in open syllables surrounded by stop consonants. As for the contact of non-identical vowels, the acoustic parameters that indicate the presence of a syllabic cluster are those established in [7], even if the vowels /i u/ do not appear in the group.

c) Monophthongization. Both vowel deletion and coalescence are considered as phonetic manifestations of monophthongizations, when only a vowel could be observed in the spectrographic displays. The sequence of identical vowels is replaced by another vowel having the same vowel quality. When vowels are not identical, the phonetic result is varying: one of the vowels initially forming the group, or a vowel with formant frequencies corresponding to a fusion of vowels.

The most common type of phonological resolution that occurs in the corpus is the hiatus, irrespective of the speaker, as shown in Table 1

|           |   | <i>hiatus</i> | <i>reduction</i> | <i>monophth.</i> |
|-----------|---|---------------|------------------|------------------|
| <i>Ag</i> | n | 28            | 25               | 20               |
| <i>Bm</i> | n | 31            | 26               | 16               |
| <i>Rm</i> | n | 38            | 28               | 4                |
| <i>Vi</i> | n | 35            | 21               | 24               |
| <i>Ra</i> | n | 35            | 16               | 19               |

**Table 1.** Number of cases of the phonological processes per each speaker (n).

|           |   | <i>hiatus</i> | <i>reduction</i> | <i>monophth.</i> |
|-----------|---|---------------|------------------|------------------|
| <i>aa</i> | n | 3             | 11               | 14               |
| <i>áa</i> | n | 13            | 7                | 1                |
| <i>áa</i> | n | 7             | 21               | 6                |
| <i>áá</i> | n | 18            | 1                |                  |
| <i>ae</i> | n | 10            | 18               | 29               |
| <i>aé</i> | n | 22            |                  |                  |
| <i>áe</i> | n | 5             | 15               | 4                |
| <i>áé</i> | n | 22            | 1                |                  |
| <i>ai</i> | n | 11            | 36               | 29               |
| <i>aí</i> | n | 20            |                  |                  |
| <i>ái</i> | n | 15            | 6                |                  |
| <i>ái</i> | n | 21            |                  |                  |

**Table 2.** Number of cases per each phonological process in each vowel combination (n).

Of the 366 vowel sequences in the data, 167 were parsed into two separate syllables (45% of the total); 116 formed diphthongs (32%) and 83 were monophthongised (23%). Differences due to speakers' preferences are found but with respect to diphthongisation and monophthongisation processes. For instance, the speaker Rm rarely uses the monophthongization (6% of the total), and inversely, the speaker Vi shows a tendency to resolve vowel sequences by means of this process (30%).

A close inspection to the data obtained from the analysis is offered in Table 2, where the frequency of occurrence of the categories is observed separately for each vowel combination in each stress pattern. All the sequences of the corpus present cases of hiatus, while reduction is not observed for [aé] [aí] [ái], and monophthongization is not observed for [áá] [aé] [áé] [aí] [ái] [ái]. As expected, if one of the vowels bears stress the group is more resistant to the application of phonetic reductions. If we pool the data of the vowel sequences attending to the stress pattern ('VV, V'V, 'V'V, VV), we observe that the pattern VV shows the highest degree of monophthongization (45%), the pattern 'VV shows the highest degree of reduction (49%) while the patterns V'V and 'V'V show a strong tendency to parse into two separate syllables the group (87% and 91%, respectively). Non surprisingly, the entirely unstressed vowel combination behaves in a different way with respect to the application of phonological processes compared to the rest of combinations. On the other hand, a difference is found in vowel sequences when the stress appears in the word-final vowel compared to vowel sequences in which the lexical stress is on the word-initial vowel. As for the first, diphthongs are more frequent than hiatuses, while for the second, there is a strong tendency to parse the group into two separate syllables.

The results do not support the claim that adjacent identical vowels always manifest as a single vowel. In all items of the corpus, regardless of the identity of the vowels, the most frequent resolution is the hiatus: 50% of the sequences of identical vowels was maintained as two segments, a percentage even higher than those found for the sequences where /i/ (46%), and /e/ appeared (48%).

In order to investigate if the position in the intonational phrase (IP) affects the phonological resolution of vowel sequences, the data were organized accordingly. Results show similar percentages of occurrence of hiatus, reduction and monophthongization, irrespective of the position in the IP. With respect to hiatus: initial=49% of the total occurrences, medial=40% and final=47%. As for the reduction cases: initial=30% of the total occurrences, medial=34%; final=30%. And for the monophthongized items: initial=21% of the total, medial= 25%; final= 22%.

### 3.2. DURATION VARIABILITY

The duration values of the vowel sequences were processed with the statistical package SPSS. The results show clear differences in the duration data of the target vowel sequences in correlation with the phonological resolution adopted. As expected from the definition of each category (hiatus = two vowels in two syllables; reduction = two vowels in a single syllable; monophthongization = one vowel), a gradual decrease of duration in the phonetic outputs of the items is observed: vowels in hiatus are longer ( $x = 171$ ,  $s.d. = 37$ ) than reduced sequences ( $x = 102$ ,  $s.d. = 23$ ) than monophthongized results ( $x = 64$ ,  $s.d. = 14$ ). An analysis of variance (ANOVA) was done on the duration data with the grouping factor 'phonological resolution' (hiatus, reduction, monophthongization), which turned out to be significant ( $F = 397,7$ ,  $p < 0.000$ ). A post-hoc Scheffe-test identified differences between each pair ( $p < 0.000$ , in all comparisons). Consequently, the investigation of the factors affecting vowel sequences must be done for hiatus, reduction and deletion separately.

#### 3.2.1. Effect of vowel quality

In order to investigate if identical vowel sequences behave differently from non-identical vowel sequences, and within the last group, if the appearance of a high vowel makes any difference, only unstressed combinations were processed.

The results do not show any apparent systematic pattern in the temporal domain, when we only pay attention to the vowels implied in the group, as in Table 3. A gradation non-high > identical > high is found for hiatus; while a gradation high > non-high > identical appears for reduced cases; and a gradation identical > high > non-high, in the case of monophthongized results. In order to determine if these differences are significant, an ANOVA with the grouping factor 'vowel quality' (identical, high, non-high) was done separately for each type. The differences were significant for the sample of reduced sequences ( $F = 11$ ,  $p < 0.000$ ) and for monophthongs ( $F = 15$ ,  $p < 0.000$ ), but not in the case of the sequences in hiatus ( $F = 0.19$ ,  $p > 0.5$ ).

#### 3.2.2. Effect of the stress

As already expected, stress pattern is an important factor in the duration of vowel sequences.

|           | hiatus |     |      | reduction |      | monophth |    |
|-----------|--------|-----|------|-----------|------|----------|----|
|           | n      | x   | sd n | x         | sd n | x        | sd |
| identical | 3      | 160 | 11   | 113       | 1814 | 78       | 15 |
| high      | 11     | 148 | 3136 | 90        | 1029 | 60       | 9  |
| non-high  | 10     | 161 | 4818 | 92        | 2029 | 58       | 12 |

**Table 3.** Number of cases (n), mean (x) and standard deviation (sd) of duration data of vowel combinations wrt to its phonological resolution.

|     | hiatus |     |      | reduction |      | monophth |    |
|-----|--------|-----|------|-----------|------|----------|----|
|     | n      | x   | sdn  | x         | sdn  | x        | sd |
| VV  | 30     | 160 | 2738 | 108       | 2510 | 69       | 14 |
| V'V | 55     | 161 | 287  | 125       | 331  | 118      | ,  |
| 'VV | 58     | 192 | 416  | 128       | 28   | ,        | ,  |
| VV  | 24     | 154 | 3665 | 94        | 1772 | 63       | 14 |

**Table 4.** Number of cases (n), mean (x) and standard deviation (sd) of duration data of the vowel sequences in different stress patterns wrt to its phonological resolution.

|            | hiatus |     |      | reduction |      | monoph |    |
|------------|--------|-----|------|-----------|------|--------|----|
|            | n      | x   | sdn  | x         | sdn  | x      | sd |
| Initial-IP | 61     | 171 | 4538 | 98        | 2026 | 65     | 14 |
| Medial-IP  | 50     | 168 | 3142 | 98        | 2231 | 67     | 16 |
| Final-IP   | 56     | 173 | 3336 | 112       | 2726 | 61     | 14 |

**Table 5.** Number of cases (n), mean (x) and standard deviation (sd) of duration data of the vowel sequences in different IP positions wrt to its phonological resolution.

In particular, a systematic decrease of duration when going from 'V'V > V'V > 'VV > VV, is found in the data, regardless of the phonetic output (Table 4). An ANOVA analysis with the grouping factor 'stress pattern' and the grouping factor 'phonological resolution' has been applied to the data. Significant differences were found due to both the stress pattern ( $F = 10$ ,  $p < 0.001$ ) and the type of phonetic result ( $F = 81$ ,  $p < 0.001$ ). An Scheffe-test showed differences in all possible pairs of stress patterns ( $p < 0.001$ , in all the comparisons).

#### 3.2.3. Effect of the position in the Intonational Phrase

The well-known effect of final lengthening is observed in the data, when facing with hiatuses and diphthongs, but not in the case of monophthongs. On the other hand, sequences in initial positions tend to be longer than sequences in medial positions, but differences are not robust. An ANOVA analysis with the grouping factor 'position in the Intonational Phrase' (Initial, Medial, Final) and the grouping factor 'phonological resolution' (hiatus, reduction, monophthongization) was applied to the duration data. Significant differences were found due to the type of phonetic result ( $F = 394$ ,  $p < 0.001$ ) but not due to the position in the IP ( $F = 0.7$ ,  $p > 0.1$ ).

## 4. DISCUSSION AND CONCLUSIONS

Differently from vocalic sequences inside words, where differences between phonetic diphthongs and phonological diphthongs are found, in the adjacency of vowels at the post-lexical level, differences between the phonetic output

of each word pronounced in isolation, and the phonetic output of the sequence, can be observed, at least, as far as word-final and word-initial vowels are concerned. In this experiment, the phonetic results of the vowel adjacency across word boundaries have been classified into three categories, which show robust differences in duration values: hiatus, reduction, and monophthongization. One way to deal with vowel adjacency is to leave the sequence unaltered and parse the vowels into separate vowels. This process has been arisen as the most frequent in the corpus. Our results show similar percentages of occurrence of hiatus, irrespective of the identity of the vowels, and the position in the IP. Moreover, although the absence of lexical stress favors the recategorization of the group as a diphthong, the probability for a hiatus to appear is still high. Our results do not support the claim in [5] about the necessary correlation between heterosyllabification and stress, nor the claim that adjacent identical vowels always manifest as a single vowel ([1], [4]).

Another strategies included syllabifying the two vowels into a single syllable, or deleting one of the intervening vowels. As for the first, our results do not allow to differentiate diphthong formation in those clusters which contain underlyingly high vowels and those that do not, as done in [4]. On the contrary, our categorization is in the line suggested by [5], where diphthong formation and glide formation are included in the same category for Northern New Mexican Spanish. As for the second —deletion cases—, the durational patterns for vowels resulting from the adjacency of two identical vowels do not differ from those proceeding from non-identical vowel sequences. Consequently, coalescence behaves in the temporal domain as vowel deletion.

The fact that results with respect to reduction and monophthongization processes are not so high as expected, can be explained by the experimental procedure adopted. It is worth indicating that the experiment reported here focuses on the combination of two content words with homogeneous consonantal context excerpted from texts reading. On the contrary, the resolution of vowel adjacency has been often observed only considering speech rate, and including all the clusters found at word boundaries, with the dangerous effect of incorporating a large amount of factors that influences duration. At this respect, it is quite natural for diphthong formation and monophthongization to be the most frequent process when a function word is implied, because function words preceding content words form a single phonological word ([6]). Instead, percentages of heterosyllabification increase when two content words are in contact, because they form two PW, and restructuration is optional. The available strategies to maintain the hiatus found in the corpus are similar to those described for other languages ([6]): separation by means of a tonal movement, insertion of a pause, lengthening of the word-final vowel, appearance of creaky voice between the two vowels forming the group.

We can conclude that the duration variability in vowel sequences across word boundaries cannot be explained only on basis of segmental or prosodic parameters, but it is necessary to incorporate the phonological categories into the analysis.

An interesting question is raised, although not treated here, with the presence of vowels that do not form part of the phonemic inventory of Spanish. The question remains as to whether or not these vowels should be considered allophones. The same can be said for diphthongs. Traditionally in Spanish only vocalic groups formed by a high vowel are considered diphthongs. Nevertheless, at least in word boundaries, other type of diphthongs (if we mean by diphthong any monosyllabic vowel group) should be taken into account. Usually the vowel that loses the property of nuclearity is high, but a non-high vowel can as well be phonetically realized as non-syllabic. In these cases, although a glide formation can occur, where the vowel /e/ changes to a semivowel, also a modification of formant frequencies in other directions are found. From our viewpoint, the same syllabic structure than for diphthongs should be hypothesized when the non-syllabic vowel is /a e o/.

## 5. ACKNOWLEDGMENTS

This research was partially supported by project grant TIC-2000-1681-C02-02 from the Spanish Ministry of Science and Technology.

## REFERENCES

- [1] T. Navarro-Tomás, T. *Manual de pronunciación española*, CSIC, Madrid, 1918
- [2] J. W. Harris, *Spanish Phonology*, Cambridge, MIT Press, 1969
- [3] J. I. Hualde and I. Chitoran, "Explaining the distribution of hiatus in Spanish and Romanian", *ICPhS 2003*, Barcelona.
- [4] M. J. Canellada and J. K. Madsen, *Pronunciación del español. Lengua hablada y literaria*, Castalia, Madrid, 1987
- [5] D. L. Jenkins, *Hiatus resolution in Spanish: Phonetic aspects and phonological implications from Northern New Mexican data*, PhD thesis, New Mexico University, 1999
- [6] M. Nespór and I. Vogel, *Prosodic phonology*, Dordrecht, Foris, 1986
- [7] L. Aguilar, "Hiatus and diphthong: Acoustic cues and speech situation differences", *Speech Communication*, vol. 28, pp. 57-74, 1999