



Universidad Nacional Autónoma de México

The study of the diphthongs of DIME corpus and its potential application for speech recognition.



Fernanda López Escobedo

El Proyecto DIME, DCC-IIMAS, UNAM 2006

CONTENTS

- * The definition of diphthong, synaloepha, hiatus and hiatus diphthongation.
- * Two visions: Trubetzky and Alarcos
- * Analysis of data
- * Partial conclusion
- * The proposal
- * General conclusion

El Proyecto DIME, DCC-IIMAS, UNAM 2006

Proposal

A new computational phonetic transcription for the diphthongs of a system of speech recognition.

Problem: to delimit the borders between the vowels which conform a diphthong.

Solution until now: to cut by half the segment of the diphthong and to assign a label for each vowel.

Proposal: to label the diphthongs as a unit.

.sil	a	s	j	e	s	l	t	r	a	v	j	e	n	.sil	o
															x
.sil	a	s	je	s	l	t	r	a	v	je	n	.sil			o
															x

El Proyecto DIME, DCC-IIMAS, UNAM 2006

Diphthongs

Diphthong: "The vowels *i*, *u*, combined to each other or accompanied each one by another vowel, within a same word, form the phonetic group that is called diphthong. The *i* and the *u* are pronounced like semivowels, when they go to the end of the diphthong, and like semiconsonants, when they go in the beginning" (Navarro Tomás, 1918:65).

Classification: a) rising: /i, u/ + /e, a, o/
b) falling: /e, a, o/ + /i, u/

El Proyecto DIME, DCC-IIMAS, UNAM 2006

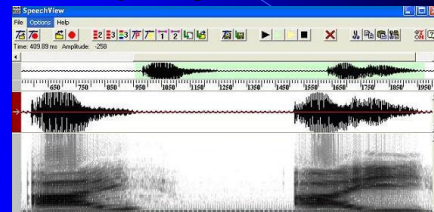
Diphthongs

Rising Diphthongs	Falling Diphthongs
ja	aj
je	ej
jo	oj
wa	aw
we	ew
wo	ow
ju	
wi	

El Proyecto DIME, DCC-IIMAS, UNAM 2006

Hiatus: "it happens a hiatus when two or more consecutive vowels are pronounced without forming a unique syllable" (Gili Gaya, 1957:117).

Diphthongs and hiatus



Diphthong: the transition of the formants is slow and its long duration. The glide is lost its own autonomy, has been subsumed by the vowel which forms the syllabic nucleus.

Hiatus: the transition of the formants is fast and has a brief duration. Both vowels are syllabic nucleus, that is why they conserve his autonomy and, therefore, spectographically those two vowels maintain their own formants well differentiated.

El Proyecto DIME, DCC-IIMAS, UNAM 2006

Synaloepha and Hiatus diphthongation

Synaloepha: "group of vowels formed by the connection of words and pronouncing in one syllable" (Navarro Tomás, 1918:69).

Example: *todo un invierno*

[t_cto Dow nim b_cjEr(no) it forms the dipthong [ow]

Hiatus diphthongation: the tendency to reduce the hiatus to diphthong.

[te.á.tro] → [tjá.tro]

El Proyecto DIME, DCC-IMAS, UNAM 2006

Two visions: Trubetzky and Alarcos

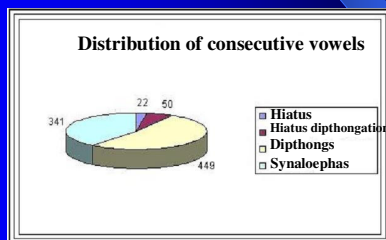
Trubetzky	Alarcos
Monophonematis: groups of sounds whose components, in the considered language, are not distributed in two syllables, are produced by an only articulatory movement and whose duration does not exceed the normal duration of the sounds.	Biphonematis: the components of the spanish diphthongs are diverse realization of the vocal phonemes, they are only combinations of the five vocal phonemes with another element.

The glide loses its autonomy and it is submit by the vowel who forms the syllabic nucleus. The formant of the syllabic nucleus becomes a continuation of the glide, in the case of rising diphthongs, and to the inverse in the case of the falling diphthongs.

El Proyecto DIME, DCC-IMAS, UNAM 2006

Analysis of data

5 dialogues of the Corpus DIME were analyzed: s2, s4, s5, s6 and s8



2006

Analysis of data

Diphthong	Quantity	Percentage
we	98	21.82 %
ej	91	20.26 %
ja	31	6.90 %
aj	19	4.23 %
je	17	3.78 %
wa	16	3.56 %
wi	4	0.89 %
oj	1	0.22 %

Numbers that represent the combinations of vowels in dipthong in the Corpus DIME.

They do not appear the diphthongs [ju], [aw], [ew], [ow], [wo].

El Proyecto DIME, DCC-IMAS, UNAM 2006

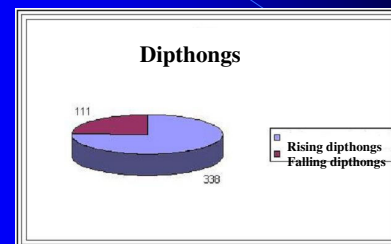
Analysis of data

Vowel	Alarcos Lerach (1965)	Navarro Tomás (1946)	Zipf y Rogers (1939)	Guirre J.y García (1993)	Perez (2003)	Cuétara (2004)	Average
/i/	8.60	4.76	4.20	6.59	7.46	7.18	6.46
/e/	12.60	11.75	12.20	14.99	14.13	15.35	13.50
/a/	13.70	13.00	14.06	13.27	12.31	14.05	13.39
/o/	10.30	8.90	9.32	10.75	9.28	9.89	9.74
/u/	2.10	1.92	1.76	2.79	3.05	3.22	2.47

Percentage comparison of the frequency of the Spanish phonemes

El Proyecto DIME, DCC-IMAS, UNAM 2006

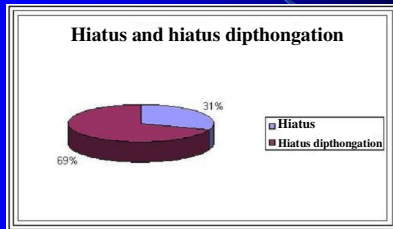
Analysis of data



"The rising diphthongs are much more frequent in Spanish than the falling diphthongs" (Alarcos, 1950:48)

El Proyecto DIME, DCC-IMAS, UNAM 2006

Analysis of data



Frequency presence of hiatus and hiatus diphthong

El Proyecto DIME, DCC-IIMAS, UNAM 2006

Partial conclusion

The diphthongs and the synaloephas are frequent in the Corpus DIME.

The hiatus and the hiatus diphthong are little frequent phenomena in the Corpus DIME.

For the case of the diphthongs and the synaloepha a new computational phonetic transcription could be very useful.

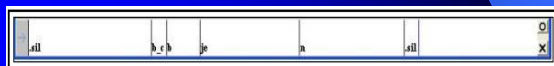
The case of the hiatus and the hiatus diphthong are part of another study.

El Proyecto DIME, DCC-IIMAS, UNAM 2006

Proposal for tagging diphthong



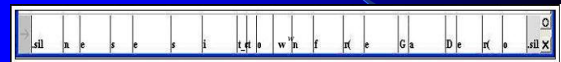
Actual computational phonetic transcription of the diphthong



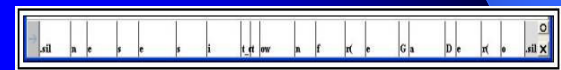
New computational phonetic transcription of the diphthong

El Proyecto DIME, DCC-IIMAS, UNAM 2006

Proposal for the synaloepha



Actual computational phonetic transcription of the synaloepha



New computational phonetic transcription of the synaloepha

El Proyecto DIME, DCC-IIMAS, UNAM 2006

The new acoustic models for the diphthong and for the synaloepha

37 acoustic models plus

ja	aj
je	ej
jo	oj
wa	aw
we	ew
wo	ow
ju	
wi	

El Proyecto DIME, DCC-IIMAS, UNAM 2006

The new pronunciation dictionary for the diphthong

Actual pronunciation dictionary	New pronunciation dictionary
b_c b j e n (good)	b_c b j e n
t_c t j e n e s (have)	t_c t j e n e s
m w e V e s (move)	m w e V e s
t_c t r(a j m e (bring me)	t_c t r(a j m e

El Proyecto DIME, DCC-IIMAS, UNAM 2006

The new pronuntiation dictionary for the synaloepha

Actual pronuntiation dictionary	New pronuntiation dictionary
a s i (that)	a s j e s t c t a
e s t c t a (is)	
n e s e s i t c t o (I need)	n e s e s j t c t o w n
u n (a)	
a o r (a now)	a o r (a w u n
u n (a)	

El Proyecto DIME, DCC-IIMAS, UNAM 2006

The new language models for the synaloepha

Actual language models	Actual language models
u n (a)	f r (e G a D e r (o (sink)
	m w e V l e (furniture)
	l a V a t c t r (a s t c t e s (dishwasher)

New language models	New language models
n e s e s j t c t o w n	f r (e G a D e r (o (sink)
	m w e V l e (furniture)
	l a V a t c t r (a s t c t e s (dishwasher)

El Proyecto DIME, DCC-IIMAS, UNAM 2006

General conclusion

Advantages	Disadvantages
Saving time in the labeling work	Increase in the acoustic models
A potential improvement in the speech recognition	Extension in the pronuntiation dictionary
	Extension in the new languages models

El Proyecto DIME, DCC-IIMAS, UNAM 2006

Thanks

El Proyecto DIME, DCC-IIMAS, UNAM 2006