Papers and Studies in Contrastive Linguistics 34, 1998 pp. 27-51

SOME REFLECTIONS ON THE NOTION OF DIPHTHONG

FERNANDO SÁNCHEZ MIRET
University of Salamanca, Spain

> "If long vowels produce methodological headaches, diphthongs are a positive migraine." (Lass 1984: 95).

1. Introduction

The notion of 'diphthong' is not well defined in the linguistic literature (Lass 1984: 95; Maddieson 1984: 161).[1] If linguists try to make an absolute definition they are confronted with the following dichotomy: are diphthongs single sounds or sequences of two sounds? This becomes evident if we compare the following definition by Catford (1977: 215): "A diphthong may be defined as a sequence of two perceptually different vowel sounds within one and the same syllable" with the one by Ladefoged (1982[2]: 171): "[diphthongs are] single vowels with continuously changing qualities."

This conflict between unity (single vowel) and duality (sequence) is present in all discussions of diphthongs.[2] In this sense, a well-known field of controversy is the question of mono- or biphonematicity which has been a favorite topic in the structuralistic phonology and in different works on experimental phonology[3]

---

[1] A preliminary version of this paper was presented at the *30th Poznań Linguistic Meeting* (1-3 May 1997). I am very grateful to Hans Christian Luschützky and Jesús Fernández González, who made some valuable comments to a second version of the paper. As usual, the responsibility for any errors rests with me. My attendance at the *30th Poznań Linguistic Meeting* was supported in part by the Spanish Ministry of Education.

[2] Greek *δι φθογγος* 'double sound' points to the duality, so that Peeters (1991: 213) considers that it is inadequate for the diphthongs /ai, au/ in the different Germanic languages. Such diphthongs are in his opinion "inherently dynamic monophthongs".

[3] The problem is unavoidable in every practical analysis. Nevertheless, Rischel (1991: 241-242) expresses a pessimistic opinion about the possibility of solving it. On the one hand, he thinks that this question is not interesting, because the decision is nearly always arbitrary. On the other hand, the internal representation of one speaker must not be necessarily coincidental with the representation of the rest of the speakers (a fact

(Slama-Cazacu 1958; Cowan et al. 1985; Berg 1986; Bertinetto 1988; Bertinetto et al. 1994). From this point of view, diphthongs, as well as long vowels, affricates, geminate consonants, and prenasalized stops are very problematic entities (see the epithets *suspicious sounds*, Pike 1947a: 251; or *"suspect" complex phonetic events*, Maddieson 1984: 161).[4] This indeterminacy arising in analyses of concrete languages (see, for example, Pike 1947b) points to a fundamental characteristic of diphthongs from the point of view of a theory of language universals, namely its Janus-like character: diphthongs share characteristics of both sequences and single segments, and this fact constitutes its essential nature.

Besides, there is another great amount of debate about the different types of diphthongs and their legitimacy to be so called. There is for example little agreement whether entities like /je, wo/ (rising diphthongs) in Romance languages should be classified as authentic diphthongs. In this issue different factors are involved:

a) the exact phonetic nature of what is transcribed with /j, w/[5];
b) the traditional idea that authentic diphthongs are always falling (like [aɪ̯], see below section 2);
c) the mostly unavoidable "English-point-of-view", since for this language the similar sequence /ju/ is generally analyzed as consonant + vowel (CV)[6];
d) the historical factors play an important role also at this point, as already Sievers (1901[5]: § 422) pointed out that: in the Romance languages the sequences /je, wo, we/ are called "rising diphthongs", since they historically derive from single vowels. Moreover, the process of diphthongization of Latin *ŏ*, *ĕ* created in some Romance languages alternations *diphthong* ~ *monophthong* (e.g., Spanish *vuelo* 'I fly, flight' ~ *volar* 'to fly'), that reinforce the diphthongal analysis of these sequences.

For the reasons I have mentioned so far, an absolute definition seems to be impracticable. In response to this problem, many linguists tend to classify diphthongs or potential diphthongs into different dichotomies. I will refer only to two examples:

1) Andersen (1972: 18) differentiates between 'segmental diphthong' ("a single segment whose central phase is acoustically heterogeneous in its temporal development, rather than presenting a steady state") and 'sequential diphthong' ("a sequence of segments, usually forming part of the same syllable").

that underlies the possibility of abduction, cf. Andersen 1973).

Furthermore, Blumstein (1991) poses the question whether phonetics can help to the solution of this kind of problems, like the mono- or biphonematic analysis of some entities. In her opinion, these questions are exclusively phonological and phonetics has very little to say.

[4] For an extensive analysis of the affricates-problem see Luschützky (1992).

[5] "Remember that there is nothing sacred about the phonetic value of a symbol" (Ladefoged 1982[2]: 36).

[6] For a different analysis see Ladefoged (1982[2]: 77-78).

2) Lehiste and Peterson (1961: 277) include under the general term of "complex nuclei" two types of entities: 'glides' and 'diphthongs'. "A *glide* is a vocalic syllable nucleus consisting of one target position, with associated formant transitions to the target, and formant transitions from the target. [...] A *diphthong* is a vocalic syllable nucleus containing two target positions."[7]

The idea behind such classifications, as shown in Table 1, is that some diphthongs behave like units (or show more unitary features) and some others behave like sequences (or show more dualistic features).

Table 1. Dichotomous classification of diphthongs.

units	sequences
segmental diphthong: a single segment	*sequential diphthong*: a sequence of segments
glide: one target position	*diphthong*: two target positions

2. "True" and "false" diphthongs

One finds from the beginning of the investigation about diphthongs some kind of idea that there should be good and bad examples of diphthongs. Two terms have often been used: 'true' and 'false' diphthong (*echt* vs. *unecht* Sievers 1901[5]: § 418; *eigentlich* vs. *uneigentlich* Jespersen 1897-99 [1904]: § 211). From this perspective some diphthongs are considered to be good representatives of the ideal diphthong, and are said to be true members of that category, while other diphthongs are viewed either as another type of entity or as bad examples of the category. I shall show only a few examples.

Sievers (1901[5]: § 418) considers true diphthongs instances like [aɪ̯, eɪ̯, aʊ̯, oʊ̯], which give an impression of unity ("[...] beide Teile für das Ohr mehr zu einer Art glatt verlaufender Einheit zusammenschmelzen"). On the other hand, Sievers considers that [iə̯, uə̯, yə̯] are false diphthongs, since they are more like sequences ("die beiden Glieder mehr selbständig und unvermittelt neben einander zu stehen scheinen").

Another use of this dichotomy affects the division between 'falling' (or 'offgliding') and 'rising' (or 'ongliding') diphthongs. The traditional idea is that only the falling ones are true diphthongs (cf. Sievers 1901[5]: § 422, who does not

[7] For Lehiste and Peterson (1961: 272) 'target position' means: "The time interval within the syllable nucleus where the formants are parallel to the time axis". Another more frequent term for this is 'steady-state'. In other cases 'target' has the meaning of canonical form of a phoneme (cf. Keating 1988: 6).

Lehiste and Peterson (1961) use the term 'glide' in a particular form for designating one kind of complex nuclei. Further, 'glide' has also in the literature two other more usual meanings. On one hand, 'glide' can refer to the transition between two steady-states or to the transition *from* a steady-state or *to* a steady-state (cf. Peeters 1991: 302). For these entities we also have the term 'transition', which I shall prefer here. On the other hand, 'glide' designates one of the components of the diphthong as opposed to the 'nucleus' (see below section 5.4) and in this sense will be used here.

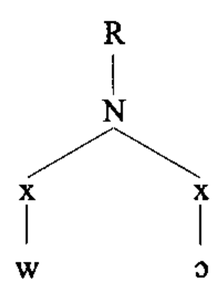
share this common opinion; Jespersen 1897-99 [1904]: § 211). An explicit argumentation for this claim is made by Donegan (1978 [1985]: 190), who takes into account prosodic arguments like the equivalence between falling diphthongs and long vowels (both count as two moras vs. rising diphthongs, which usually count only as one mora), or the different role of falling and rising diphthongs in rhyme as in the examples in: both parts of falling diphthongs count in rhyme, but the glide of rising diphthongs does not count.

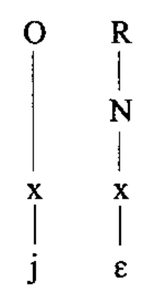
paid [pɛi̯d] : *raid* [rɛi̯d] rhyme

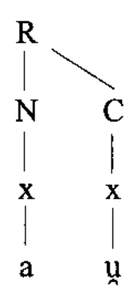
paid [pɛi̯d] ≠ *red* [rɛd] do not rhyme

feud [fi̯ud] : *mood* [mud] rhyme

In her analysis of Italian diphthongs, Marotta (1988) distinguishes three kinds of tautosyllabic vocalic sequences, of which only one can be considered a "true" diphthong. Marotta divides these Italian vocalic sequences as follows: a) the rising diphthong /wɔ/ (e.g., *uomo* 'man'), that behaves as a vocalic nucleus; b) the rising diphthong /jɛ/ and the other diphthongs with /j/ (e.g., *ieri* 'yesterday', *piano* 'slowly', *fiume* 'river'...), that behave like a CV sequence; c) the falling diphthongs (e.g., *mai* 'nunca', *poi* 'afterwards', *sei* 'six', *pausa* 'break', *Europa* 'Europe'), that are analyzed as divided between nucleus and coda.[8]

a)


b)


c)


For Marotta (1988: 407) only diphthongs that are entirely dominated by the node nucleus on the representation of the syllable can be considered as "true" diphthongs. Consequently, the only "true" diphthong in Italian would be the rising /wɔ/.[9]

Another such case of dichotic classification can be found for Dutch diphthongs. This language is said to have the 'genuine diphthongs' /ɛi, ʌy, au/, and the 'pseudo-diphthongs' /aj, oj, uj, ew, iw/. The genuine ones are more like unitary segments, while the pseudo-diphthongs act like sequences of vowel + consonant (VC) (Mioni 1973: 378-379). Phonetically the genuine diphthongs seem to be the

[8] .Some of these Italian falling diphthongs are not easily differentiated from hiatuses (cf. Marotta 1987).

[9] Through its layered formalization Autosegmental Phonology offers a representation of the conflict between unity and duality in diphthongs and in other complex entities (the same applies for affricates, cf. Luschützky 1992: 136). One finds a hint of this idea in Pike (1947b: 158).

product of a single gesture, while the pseudo-diphthongs seem to be the result of two gestures (Collier et al. 1982).

Finally, as I have said, Lehiste and Peterson (1961) distinguish in American English two types of complex nuclei.[10] On the one hand, they refer to [ɑi̯, ɑu̯, ɔi̯] (e.g., *tight, loud, voice*), which have two target positions and can be called 'diphthongs'. On the other hand, they find that [ei̯, ou̯, ɜ] (e.g., *fate, lope, hurt*) contain a single target position, "and therefore should not properly be classed as diphthongs" (Lehiste and Peterson 1961: 275).

I summarize the preceding sketched classifications in Table 2.

Table 2. Classifications of 'true' and 'false' diphthongs.

	criterion	true diphthongs	false diphthongs
Sievers	auditory impression	unity [ai̯, ei̯, au̯]	sequence [iə̯, uə̯]
Jespersen	tradition	falling	rising
Donegan	prosody	falling	rising
Marotta	syllabic structure	branching nuclei [wɔ]	onset + nucleus [jɛ], nucleus + coda [au̯]
Dutch diphthongs	phonotactics, articulation	VV /ɛi, ʌy, au/	VC /aj, oj, uj, ew, iw/
Lehiste and Peterson	acoustics	diphthongs [ɑi̯, ɑu̯, ɔi̯]	glides [ei̯, ou̯, ɜ˞]

As can be inferred from Table 2, in each case the distinction has been made from a different point of view. As a consequence, contradictions arise. For example, if for Jespersen and Donegan falling diphthongs are the best representatives of the category, in the analysis of Marotta a rising diphthong is the only one that deserves this label.

Instead of viewing the problem in terms of "true" and "false" diphthongs, I would like to propose a dynamic interpretation based on some principles of Natural and Cognitive Phonology. I use 'dynamic' in the sense that one can better understand an object if one knows the rank of its possible origins and the plausible changes it may subsequently suffer: the phonological processes that create, change and eliminate diphthongs. My claim is that we can consider the diphthong a prototypical category governed by such processes. From this point of view some diphthongs are more prototypical than others, or they match better than others some of the characteristics of the category. In addition to this, the processes of diphthongization and assimilation can explain the prototypical features of the category and the relationship between their members. I think that this approach to the problem allows one to see the variety of facts from one single point of view, as well as the synchronic and diachronic dynamic aspects of these sounds.

[10] See a similar analysis in Pike (1947b).

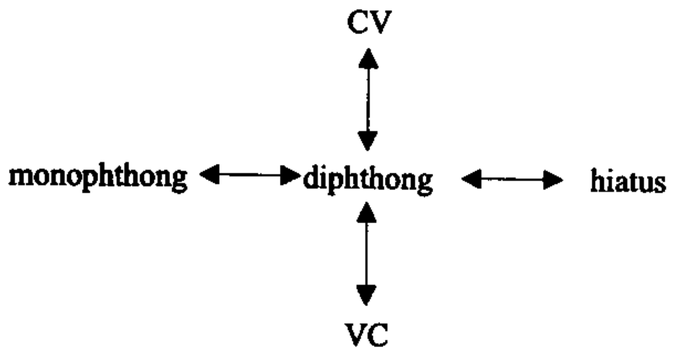
I will also argue, in the vein of the studies of typology and universals (Crothers 1978; Maddieson 1984; 1991), with data collected from different phonological inventories.

3. Diphthongs and other categories

We need to distinguish between diphthongs and other vocalic entities, which are placed at two opposite points in the scale of unity/duality, namely the monophthong and the hiatus. This distinction is not easy. Diphthongs are characterized by the presence of formant movement. However, it has been reported that monophthongs also exhibit some quantity of formant movement (Strange 1989; Nearey 1989; Andruski and Nearey 1992; but see the discussion by Harrington and Cassidy 1994; see also section 5.2). Likewise, there are difficulties when trying to discriminate between diphthong and hiatus in concrete analyses (Marotta 1987; see below also section 5.2).

A look at the historical evolution shows the different changes that can involve diphthongs: diphthongs can monophthongize (Latin *oe* > Proto-Romance *e*), can evolve into a hiatus (Latin *au* > Romanian *a.u*), or consonantize their pre- or post-vocalic glide (Italian *uomo* 'man' > dialectal *vomo*; Romance *ei* > Raeto-Romance *ek*). We also find changes with the opposite direction: *monophthong, hiatus, VC and CV > diphthong* (Latin *ĕ* > Spanish *ie*, Spanish *ma.íz* > dialectal *maiz*; Latin *multu* > Portuguese *muito*, Latin *plenu* > Italian *pieno*).

This suggests that diphthongs cover an area of the vocalic possibilities of human beings situated in the middle position between a monophthong and a hiatus, and also between a VC sequence and a CV sequence. This is represented below.



Diphthongs can be situated in the space delimited by these four categories and have the possibility to place themselves close to every one of these poles.

4. Data

I draw the data that will support my analysis from three different sources. The UPSID (UCLA Phonological Segment Inventory Database; Maddieson 1981, 1984 and UPSID 1992)[11], the SPhA (= Stanford Phonology Archive; Crothers et al. 1979) and Weeda (1983). In order to attain some homogeneity among the data I have made

a few adaptations of the transcriptions given in the sources.[12] The numbers in round brackets that follow some diphthongs in the tables and in the text refer to the number of languages that attest the diphthong (if no number follows, the diphthong is attested only in one language of the inventory). I present in figures (D1-D5) the collected diphthongs from the mentioned sources. The first elements of diphthongs are shown on the vertical axis; the second elements are listed on the top of the figure; the oblique line of boxes separates falling (below the line) from rising (above the line) diphthongs[13]; the dashes inside the boxes symbolize that a sequence of two equal sounds is not a diphthong. The double line in figure (D1), (D2), (D3) and (D5) separates the diphthongs involving a central vowel from the other diphthongs.

UPSID (1992) includes 451 languages. The diphthongs listed in the graph (D1) belong to 48 languages.[14] The reason why so few languages are considered to have diphthongs is that UPSID takes into account only monophonematic entities (cf. Maddieson 1984: 133).

The SPhA includes 197 languages; 55 of them have some kind of diphthongs. This inventory offers, besides the phonological system, also information about allophonic realizations and biphonematic diphthongs (however, sometimes we find only the reference to the existence of such diphthongs without concrete listing, which makes it impossible to take them into account in my computations).[15]

In the figures D2-D4 I present the data of the SPhA about diphthongs under three different status: monophonematic, allophonic and biphonematic diphthongs.[16]

---

[12] Personally I prefer transcriptions like [a̯i, i̯a], where [̯] signalizes the glide (cf. Sievers 1901[5]: § 410). In the sources one finds (among others) [a̯i, aj, ai, a^{i}]. I have respected the transcriptions of the glide and have altered only in some cases the symbols for the nucleus.

[13] Diphthongs that are placed very close to the line (e.g., [iu, ui, oe, eo]) fluctuate between falling and rising realizations.

[14] Bladon (1985: 147) says to have found 78 languages with diphthongs in Maddieson (1981). I find in Maddieson (1981) and (1984) only 23 languages with diphthongs (cf. Maddieson 1984: 133). Bladon gives a table with the data he has found: the number of languages which attest every kind of diphthong is bigger than in my figure (D1), but the general tendencies one can extract from the data are the same.

In (D1) I do not take into account the nasal, pharyngealized and breathy voiced diphthongs. Also, for reasons of space, the following attested diphthongs involving secondary vowels are not shown in the figure: [ɔy, aɯ (3), aɪ (2), ɛɯ, iʌ, æʉ, ɤi, ɵy, uʌ, yø, əɯ, ɯa (2), ɯi (2), ɯʌ, ɯə (2), ɜi, ɪa, ɨi (2), ɪə].

[15] Very often the SPhA just refers to sequences vowel + glide (VG) or glide + vowel (GV), as for instance in Portuguese, Spanish or Rumanian. All these cases are not included in the figures, since the SPhA does not provide the concrete form and number of such VG or GV sequences, which means that the number of languages with possible diphthongs in the SPhA is actually bigger than 55.

[16] In (D2-D4) I have fused SPhA's distinction between [e̯, o̯] (mid vowels) and [e, o] (mid high vowels) (this may have little consequences in the frequency of the diphthongs with mid high vowels); cf. the same policy by Maddieson (1984: 203) for the inventories of the UPSID. Somewhat in contradiction to this position, but in order to give an acceptable representation to SPhA [ee̯, oo̯] I transcribe these diphthongs in the figures as [eɛ, oɔ]. I have also excluded from the figures the nasal and retroflexed diphthongs and I do not differentiate between long and short diphthongs.

For reasons of space, I do not include in (D3) the following diphthongs [ɯ̽w, ɤw, ɤw̽, w̽ɪ, ɪ̯, ɤe, oɤ, ɔ̝, ɛ̝, ɵw, ua̯, ya̯, ja̯, ø̯a̯, w̽ə].

Figure D1. diphthongs in UPSID (1992).

	i	u	ɪ	ʊ	e	o	ɛ	ɔ	æ	a	ə
i	—	iu (2)			ie (6)	io	iɛ		iạ	ia (5)	iə
u	ui (8)	—				uo (4)				ua (3)	uə (5)
ɪ			—		ɪe						
ʊ	ʊi			—							
e	ei (7)	eu 3			—	eo (3)				ea	eə
o	oi (6)	ou 5			oe	—				oa (2)	
ɛ	ɛi (3)					ɛo	—				
ɔ	ɔi (6)							—			
æ	æi (3)								—		
a	ai (19)	au (18)			ae (3)	ao (4)	aɛ			—	
ə	əi (2)	əu (4)									

Figure D2. Monophonematic diphthongs in SPhA.

	j	w	ɪ	e	o	ɛ	ɔ	æ
j	—			je (4)	jo	jɛ (2)		jæ
w		—			wo		wɔ	
ɪ			—					
e				—				
o		ow			—			
ɛ						—		
ɔ					ɔo̯		—	
æ				æe̯				—
ɜ			ɜɪ̯					

Some other diphthongs are not included in (D3): [iạ, eɛ̯, oɔ̯] and [ɛ̯ɪ, o̯u, o̯i, ə̯ɪ, ạe, ạə]. According to principles about diphthongs most widely defended (see section 5) the above sequences can hardly be accepted as diphthongs. However, the cases [iạ, eɛ̯, oɔ̯] can be considered well-formed diphthongs under specific conditions (see section 5).

Not included in the figure (D4): [aɯ, ɑʊ, o̯i, ɔ̯y, o̯u, o̯o, o̯ɛ, ʌw, œy, ɸy, ɛɑ, ɯj].

Figure D3. Allophonic diphthongs in SPhA.

	j	w	i	u	ɪ	ʊ	e	o	ɛ	ɔ	æ	a	ə
j	—						je	jo		jɔ	jæ	ja	
w		—			wɪ			wo (2)	wɛ			wa (2)	
i	ij (4)		—										iə̯ (2)
u	uj	uw		—									uə̯
ɪ	ɪj		ɪi		—								iə̯ (2)
ʊ		ʊw		ʊu		—							ʊə̯
e	ej (7)		ei				—		eɛ				eə̯ (2)
o	oj (2)	ow (3)		ou (2)				—		oɔ			
ɛ	ɛj (2)							ɛo	—			ɛa	ɛə
ɔ										—			ɔə̯ (3)
æ											—		æə̯ (3)
a	aj (3)											—	
ɐ	ɐj												
ɜ													

Figure D4. Biphonematic diphthongs in SPhA.

	j	w	i	u	ɪ	ʊ	e	o	ɛ	ɔ	æ	a
j	—											
w		—										
i		iw (2)	—									
u	uj (3)		ui (2)	—								
ɪ					—							ɪa (2)
ʊ					ʊɪ	—						ʊa (2)
e			ei (2)		eɪ		—	eo (3)			eæ	ea
o	oj	ow	oi (2)	ou (2)	oɪ	ou	oe	—	oɛ			oa
ɛ	ɛj	ɛw	ɛi	ɛu					—			
ɔ	ɔj (2)				ɔɪ (2)					—		
æ		æw			æɪ		æe				—	
a	aj (3)	aw (3)	ai (5)	au (4)	aɪ (2)	au	ae	ao	aɛ			—

Figure D5. diphthongs in Weeda (1983).

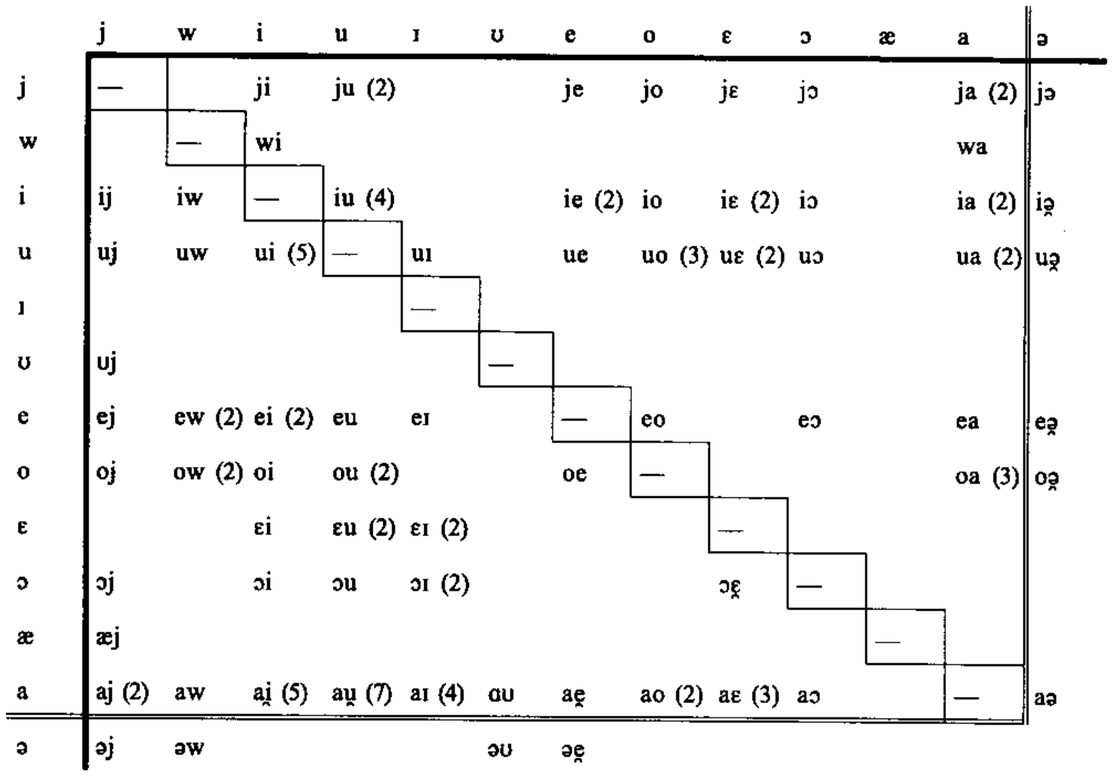


Figure (D5) reflects the data extracted from Weeda (1983), which is not an inventory of phonological systems, but a study about diphthongs. In this work one does not find any information about the (mono- or biphonematic) status of the diphthongs. The author analyzes 26 languages, of which 21 have diphthongs.[17]

Figure (D6) summarizes the data of (D1-D5) focusing on the most frequent diphthongs. I include in (D6) those diphthongs that are attested in more than one language. The diphthongs are arranged left to right from the most frequent to the less frequent.

Figure D6. Most frequent diphthongs

+				–
(D1) ai, au	ui	ei	oi, ɔi ie	ou, ia, uə, uo, ao, əu, eu, ɛi, æi, ae, ua, eo, iu, əi, oa
(D2)			je	jɛ
(D3) ej	ij	aj, ow, ɔə̯	wa, wo, oj, ou, ɛj, eə̯, ɪə̯, iə̯	

[17] Not included in (D5): [yu, yw, wɑ, ɔɑ, iɑ, ʌy, ɑɪ, ɑj (2), eɑ (2), iɨ̯, ɔɨ, əɨ, ɨj, ɨw, ɨ̯i, ɪɛ, ɪə]. I have also excluded [ə̯ɛ], which has a doubtful syllabification.

(D4)	ai	au	aw, aj, uj, eo	ui, oi, ei, ou, aɪ, iw, ɪa, ʊa, ɔj, ɔɪ	
(D5)	au	ui, ai	iu, aɪ	uo, oa, aɛ	ju, ja, ie, iɛ, ia, uɛ, ua, ew, ei, ow, ou, ɛu, ɛɪ, ɔɪ, aj, ao

5. A set of preferences for the diphthong

I would like to propose that the diphthong can be seen as a prototype category (for the applicability of Cognitive Theory in phonology see Nathan 1986, 1989, 1996; Hurch and Nathan 1996). This means that the category 'diphthong' cannot be defined by the presence or absence of some necessary and sufficient conditions of membership. Instead, it is necessary to find out a series of features that contribute in different degrees to the building up of the category. According to those features, some diphthongs match better than others the characteristics of the prototypical diphthong. This conception has the following advantages:

a) it allows a simple taxonomy to be exceeded, and the different types of diphthongs that have been discussed in the linguistic literature can be organized within one category;
b) it explains the possible evolutionary routes of a diphthong (as well as its possible origins);
c) it situates the diphthong in an intermediate point between the categories of monophthong, hiatus, VC and CV.

I shall propose several prototypical features (universal preferences) of the diphthong as criteria for the organization within the category.[18]

Arguments for the determination of a preference are:

a) the number of languages showing the pattern and
b) its possible motivation (Maddieson 1991).[19]

Natural Phonology claims that processes act as "shapers of phoneme systems" (Hurch and Nathan 1996: 241). On one hand, a group of foregrounding processes (the so called prelexical processes) are responsible for the form of the phonological systems and the phonotactics of the languages and "constitute, at least in part, universal prototypicality judgments about sounds" (Hurch and Nathan 1996: 241). On the other hand, backgrounding processes build an explanation of the form these pho-

[18] Some of these universal preferences have already been discussed (especially by Weeda 1983). Some aspects less developed by Weeda will be further investigated in this section.

[19] The frequency of some concrete types of diphthongs points to the universal preferences of the category. However, frequency does not explain anything in itself, on the contrary, it has to be explained (Dressler 1989: 111, 1996; Luschützky 1992: 133). The explanatory arguments for the frequency of one concrete type of diphthong are based in general principles as the functions of the phonology (to make languages pronounceable and perceptible). Such functions are served by the phonological processes (Dressler 1984). In the case of diphthongs two processes have a fundamental importance: diphthongization and monophthongization (assimilation).

nemes show in surface representations. I shall argue that the processes of diphthongization and its converse monophthongization or assimilation explain the patterns shown by diphthongs. I shall discuss the following questions (some of them formulated, but not discussed by Lass 1984):

5.1. Are all vowel combinations possible in diphthongs? Which are better? And why?
5.2. Is there any preference for a particular height relationship between the two members?
5.3. What is the structure of diphthongs?
5.4. Is there any preference for the position of the nucleus?

5.1. Are all vowel combinations possible in diphthongs? Which are better? And why?

The first preference of diphthongs consists of their tautosyllabicity. In this case, we are dealing with an indispensable condition.[20] Consequently, the best diphthong is that which is clearly tautosyllabic. However, not every diphthong fulfills this condition to the same degree. In some cases the linguist cannot determine if one is dealing with a tautosyllabic or a heterosyllabic entity: this is the case, for example, with the sequence [iẹ̯] in the South German dialects (cf. Sievers 1893[4]: § 392), or in Leonese (Northern Spain).[21] The same problem affects the Italian falling diphthongs (cf. Marotta 1987).

The principal factor involved in this issue is the sonority of the components of the diphthong. Every syllable has only one nucleus and, for this reason, diphthongs formed by two sounds of great sonority are disfavored, since both sounds could be nuclei and that means that they tend to form a hiatus. The data support this statement, since the combinations of the most open vowels [a, æ, ɔ, ɛ] with themselves [aɛ, aɔ, aæ...] are scarcely attested, which is no wonder, since they are not good examples of diphthongs. If such a sequence remains as a diphthong, the tendency is to increase the difference of sonority between the two members via the loss of sonority in one of them (*a.e* > *aẹ̯* > *aị̯*, e.g., a.e.re > Spanish *aire* 'air'). In the few cases where we come across such diphthongs in the inventories one expects that in the actual realization one of the elements (the glide) decreases its sonority in some way.

Besides the sonority of the elements, another decisive factor for the issue of tautosyllabicity is the organization into nucleus and glide (see below section 5.3-5.4).[22] Note that the area of conflict between tauto- and heterosyllabicity mostly

[20] As Menzerath (1941: 11) states, "in der Tat ist eine richtige Diphthongdefinition nur mit Bezug auf die Silbe möglich."

[21] In Leonese [iə] is usually described as a hiatus with a diphthongal realization (in weak contexts, polysyllabic words and words outside the phrasal accent, Krüger 1923: 22).

[22] In some other cases the indeterminacy between diphthong and hiatus is due to historical or morphological factors, as in the case of Spanish /iu, ui/ (cf. RAE 1973: § 1.4.11e, g): /ui/ is a hiatus in verb forms (*constru-imos* 'we build'), in derived forms with a suffix *-ita* (*jesuita* 'Jesuit') or in cases with conservation of the Latin accentuation (*gratu-ito* 'free of charge').

affects falling diphthongs. As I have said, this is the case in Italian. One reason can be the fact that falling diphthongs usually are longer than rising ones. Equally the glide of falling diphthongs shows more vocalic features than the glide of rising diphthongs (Avram 1975; the historical evolution confirms this point, since the consonantization of the glide occurs more often in rising than in falling diphthongs). The consequence is that from this angle, rising diphthongs could be seen as more prototypical than falling ones, since they fulfill more clearly the prerequisite of tautosyllabicity and in some analyses a rising diphthong falls completely under the nucleus, while a falling diphthong is split between the nucleus and the coda. This is the view partially taken by Marotta (1988) regarding the Italian diphthongs (cf. section 2). Such an analysis finds some support from the results of a psycholinguistic experiment carried out by Bertinetto et al. (1994): they analyze the cohesion between the two members of several vowel clusters, and find out that rising diphthongs are less separable than falling ones.[23]

Summing up, the tautosyllabicity and the cohesion of the two members play a role in the unity of the diphthong. We shall see that for the other tendency, namely the need for duality, other preferences are in place.

5.2. Is there any preference for a particular height relationship between the two members?

The components of a diphthong must show a difference in quality, otherwise there would be no case to speak about diphthong (= two sounds; observe the formation of a diphthong in Spanish *no peleé tanto* 'I didn't fight so much' > American Spanish *no pelié tanto*, Alcina and Blecua 1975: 413). These two aspects, tautosyllabicity (section 5.1) and difference in quality, are obligatory and appear in every definition of the diphthong. They form the essential characteristic of this entity, which can be formulated as unity in duality or duality in unity (it depends on what one wants to stress).

Once we have stated the necessity of a qualitative difference between the two members of a diphthong, the following question is: how big must it be? Weeda (1983: 149) and Lindblom (1986: 36) propose that those diphthongs are preferred

[23] In one of the experiments the task consists in the repetition of the first syllable, and the results are (Bertinetto et al. 1994: 23, Table 3):

diphthong type	split	cohesion
falling diphthong	*laigo* → *lalaigo* (50.8%)	*lailaigo* (45%)
rising diphthong	*ziako* → *ziziako* (12.5%)	*ziaziako* **(85.4%)**

The results for the task of substitution of the first syllable with the sequence /vu/ are:

diphthong type	split	cohesion
falling diphthong	*laigo* → *vuigo* (43.3.%)	*vugo* (52%)
rising diphthong	*ziako* → *vuako* (10.8%)	*vuko* **(75%)**

which exhibit a greater trajectory in the vowel space or, in other words, those diphthongs with the maximum perceptual differentiation of endpoints. Lindblom (1986: 36) proposes the hierarchy: /aj, aw/ » /ej, ow/ » /uj, iw/.[24]

A slightly different point of view is taken by Maddieson. Instead of focusing on the distance between the two members of the diphthong, Maddieson (1984: 134) states that "diphthongs that begin or end with a high vowel element are preferred over those which lack such an element", and he insists on the fact that it does not mean that the diphthongs with a maximum trajectory are preferred, since diphthongs as /ei, ie, ou/ are quite often attested in the languages of the world.

As I have already mentioned, if both members of a diphthong have a high degree of sonority, the tendency is to have a hiatus (section 5.1). This fact can be seen as the reason for the preference of a high vowel in diphthongs, which allows the other element to have a greater degree of sonority and also usually to be the nucleus.

Further, one can suppose that the trajectory in the vowel space must be, as a tendency, greater in the case of biphonematic diphthongs, than in those which are still allophonic. The data from the SPhA seem to confirm this tendency: on one hand /ai, au/ are the most frequent biphonematic diphthongs, cf. (D4). On the other hand, the most frequent allophonic diphthong is [e̯i], cf. (D3).[25] The data are given in Table 3.[26]

Table 3. Diphthong frequency.

diphthong type	diphthong frequency
(D4) biphonematic	ai (8) au (7) ui (5)
(D3) allophonic	ei (8) ou (5) ij (4)

The fundamental question at this point would be the notion of sufficient vs. maximal contrast between the two members of diphthongs. The studies on the perception of diphthongs have mainly concentrated on the parameters that listeners utilize for discriminating different diphthongs. For example, Bladon (1985) suggests that the spectral change which necessarily takes place in a diphthong plays two roles: on one hand it acts as a "weighting flag" which alerts the auditory system of the presence of a diphthong; on the other hand, it appears to be a pointer to regions (steady-states) of spectral significance for the recognition of what diphthong is being perceived. However, as far as I know, little attention has been paid in the phonetic studies to the amount of change necessary to pass from a monophthong to a diphthong ("threshold

[24] Lindblom (1986: 36) formulates this idea of a big trajectory in terms of the nuclei. In his opinion, "diphthongs are favoured according to the degree of sonority of their nuclei." He finds confirmation for his scale in Edström (1971).

[25] This tendency seems to be contradicted by the fact that among the biphonematic diphthongs registered in the SPhA we also find examples with a small trajectory in the vowel space, like [oe, ou̯, eo, oɛ]. But it still remains the fact that these are not the most frequent types.

[26] The numbers in (6) correspond to (D4) /aj/ + /ai/, /aw/ + /au/, /uj/ + /ui/ and (D3) [ej] + [ei], [ow] + [ou].

change", Nearey and Assmann 1986: 1305).[27] More attention has been devoted to the matter of the duration of the transition. As Peeters (1991: 313) points out, it seems that a transition longer than 80-100 ms. is needed in order to perceive a diphthong.[28] The same author argues that it is not clear to what extent non-phonetic factors (phoneme- or language-specific characteristics) impose some constraints on what is perceived as a monophthong, a diphthong or a VC sequence.

It is difficult to speak about universal minimal and maximal contrast between the members of the diphthong. However, it is feasible to obtain some prototype effects from the distribution of the different types of diphthongs in the data (Nathan 1989: 59). In this sense, the observation of Maddieson on the frequency of [j, w, i, u] in these entities is significant and is confirmed in the different data in Table 4, where the numbers correspond to tokens.

Table 4. The frequency of [j, i, w, u] in diphthongs. Token numbers.

sources	diphthongs with [j, i, w, u]	diphthongs without
(D1)	104	18
(D2)	11	3
(D3)	45	16
(D4)	37	26
(D5)	82	28

[27] Already by Sievers (1893[4]: §391): "Ein allgemeineres Abstandsminimum oder -maximum der Componenten lässt sich nicht angeben." The question seems now to be more complicated. The traditional opinion was that diphthongs and monophthongs can be distinguished by the fact that monophthongs are stable sounds along their duration (Vowel targets = canonical forms of vowels: represented articulatorily as static vocal tract shapes, acoustically as points in the acoustic space). Against this view, some phoneticians have pointed to several dynamic aspects (change over time in spectral structure), which should be relevant for the perception of monophthongs. On one hand, supposed monophthongs, such as the ones found in Canadian English /ɪ, ɛ, æ/ show significant formant change (comparable to the change of the diphthongs [eɪ̯, oʊ̯] = /e, o/). On the other hand, dynamic cues, as intrinsic duration, formant change and consonantal transitions have been shown to play a role in the vowel classification (cf. Nearey and Assmann 1986; Strange 1989; Nearey 1989).

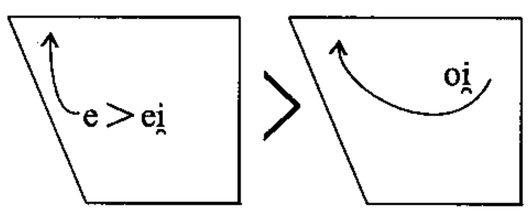
This hypothesis has been criticized by Harrington and Cassidy (1994). In their perceptual experiments the dynamic factors play a role only for the perception of diphthongs, but not monophthongs. As they point out, the proponents of the dynamic theory consider the notion of vowel target to be irrelevant, but for Harrington and Cassidy this concept is crucial in the distinction between monophthongs and diphthongs.

For the time being the only possible conclusion seems, as Harrington and Cassidy (1994: 369) say, "that the theory that vowels are inherently dynamic warrants further investigation". The different studies in experimental phonetics have revealed part of the huge complexity of parameters involved in the production and perception of sounds, but we have not yet a unitary model that could coherently articulate all these findings.

[28] This threshold is the result of different studies on Germanic languages.

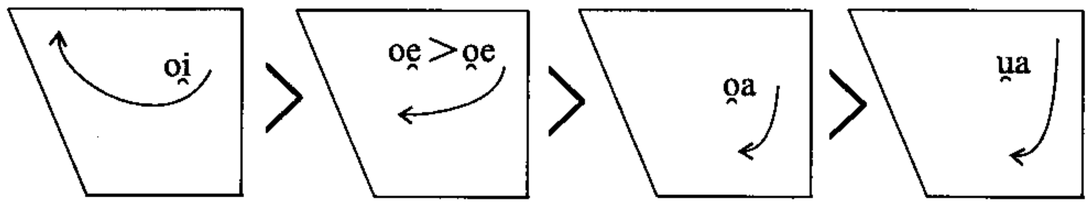
The question is what does this fact tell us and how can we explain it? I see an answer to this question in the theory of diphthongization of Stampe (1972) and Donegan (1978). The prelexical process of diphthongization defines the prototype effects for this particular class of sounds. According to these authors, the diphthongization of a vowel consists in the polarization of its two fundamental features, sonority and color, which dissimilate to the extremes of the vowel space (Stampe 1972: 584). Consequently, an optimal diphthong has a maximal sonorous nucleus and a maximal colored glide. We can observe the effect of such a process in its diachronic manifestation, for example in the evolution of Latin *ē* > French *e̯i* > *o̯i*. In the first phase we find polarization of sonority: the second part of the diphthong rises. In the second phase the polarization affects the color of the first element; see the representation in Diagram 1.

Diagram 1. Diphthongization of vowels *e* > *e̯i* > *o̯i*



At later phases of the evolution of the Ancient French diphthong [o̯i] we find the other process involved in the history of diphthongs, namely assimilation: *o̯i* > *o̯e* > *o̯e* > *o̯a* > *u̯a*; see the representation in Diagram 2.

Diagram 2. Assimilation of the diphthongs *o̯i* > *o̯e* > *o̯e* > *o̯a* > *u̯a*.



This process of assimilation probably began in less stressed contexts and then extended to all positions. Note that in the last phase the glide rises from [o̯] to [u̯]. Here we see a manifestation of the tendency towards high glides, which is active not only in the process of diphthongization of a monophthong but also during the posterior evolution of the diphthong.

In the sense of Stampe's theory of diphthongization, the high frequency of [j, w, i, u] in diphthongs is an effect of the glide's tendency to be colored. Further, Stampe's idea of an optimal diphthong as an entity which polarizes its two members finds corroboration in Figures (D1), (D3), (D4), (D5), where we can ascertain that diphthongs tend to concentrate on the most distant points from the oblique line, that is, where the distance between the two members is greatest (in terms of sonority or in terms of color, or both).

We can try to go further with this investigation of the possible combinations of vowels into diphthongs asking what is the role of the different phonetic features in this issue. The most important ways of qualitative distinction in the vowel space are:

a) difference in height;
b) difference in frontness vs. backness.

One can suppose that for the aim of producing maximal differentiation, some collaboration of these features should take place. In fact, both differences involving height and frontness vs. backness are almost always present in biphonematic diphthongs, cf. (D4). This is expected, since such diphthongs tend to be maximally differentiated. A more balanced situation is expected for the allophonic diphthongs; in this case the difference in height seems sufficient to account for a great number of diphthongs, cf. (D3). The frequencies of each type are shown in Table 5.

Table 5. The frequencies of various differentation features for two types of diphthongs: biphonematic, and allophonic.

diphthong type	total number	height + front vs. back	height	front vs. back
(D4) biphonematic	38	25	9	4
(D3) allophonic	38	23	15	

These data point to a predominance of height (vs. the feature frontness/backness), that is of differences in sonority, in the formation of diphthongs. There are a few cases where the two parts are only distinguishable by means of the opposition front vs. back, as in [iu, ui, oe, eo]. The diphthongs [ɛɔ, ɔɛ, æɑ, ɑæ…] are not attested; these hypothetical diphthongs are formed with elements of great sonority, a combination which is strongly disfavored (section 5.1). Further, it must be noted that the diphthongs formed only by means of the difference front vs. back pose the problem of differentiating between nucleus and glide (as in the case of /iu, ui/ in Spanish, RAE 1973: § 1.4.11a).

On the other hand, diphthongs containing the central vowels [ə, ɜ, ɐ] are attested in the SPhA only as monophonematic or allophonic, cf. (D2) and (D3). The most frequent central vowel in these diphthongs is [ə] and it is usually found in the glide position, cf. (D6). Furthermore diphthongs formed only with central vowels are uncommon: in the UPSID we find [əɪ, ɪa, aɪ] and in Weeda (1983) [ɪə, əɪ], but every case is exemplified only in one language. The reason for this avoidance of 'central diphthongs' can be seen in the fact that central vowels are achromatic, so that they are not very adequate for cases of polarization.

In conclusion, some distance between the members is needed for the perception of a diphthong. The greater it is, the more perceptible the diphthong is. The prototypical biphonematic diphthong has the maximal distance between its two members (as one can conclude from the data of the SPhA). The fact that the great majority of diphthongs present a high element can be explained as the consequence of two factors:

a) the process of diphthongization, which consists in the creation and increase of the differences between the two members of the diphthong;
b) the fact that two elements with a high degree of sonority tend to convert into a hiatus.

5.3. What is the structure of diphthongs?

The point I want to discuss is the internal organization of diphthongs into nucleus and glide. The nucleus is the most perceptible part and tends to coincide with the element with the greatest degree of intrinsic sonority. This is called the *sonority principle* (*Sonoritätsprinzip*: Jespersen 1897-99 [1904]: § 198) and predicts that the diphthongs which are below the oblique line in (D1-D5) are falling, and that those above this line are rising. This is a traditional idea (Luick 1891: 337; Sievers 1893[4]: § 392; more recently Donegan 1978 [1985]: 191), and this is the reason why I have excluded diphthongs like [i̯a, e̯ɛ, o̯ɔ] and [e̯ɪ, o̯u, ə̯i, ə̯ɪ, a̯e, a̯ə] from the tables (D1-D5) (see fn. 16, 17), since they imply an implausible organization. Nevertheless, cases like [i̯a, e̯ɛ, o̯ɔ] can correctly be interpreted as diphthongs. In [i̯a, e̯ɛ, o̯ɔ] the postulated glide has more intrinsic sonority than the nucleus, a fact that runs against the *sonority principle*. The same applies to diphthongs like [i̯ə, u̯ə]. As Schubiger (1977[2] [1989]: 116) points out, the glide in these cases is produced with less expiratory intensity, i.e. [a̯, ɛ̯, ɔ̯, ə̯] can correctly be less sonorous than the nuclei [i, e, o, u]. Note that these diphthongs were called "unecht" by Sievers (section 2) and they are the product of the so called 'centering' or 'ingliding' diphthongization, which mostly affects lax vowels (Sánchez Miret 1996: § 2). On the other hand, diphthongs like [e̯ɪ, o̯u, ə̯i, ə̯ɪ, a̯e, a̯ə] are actually unpronounceable, no matter how much one can diminish the expiratory intensity on the first element. This points to an important restriction on the articulatory capabilities of humans.

The *sonority principle* finds its parallel in the historical evolution in cases of syllabicity change, as in *rege* > Old French [roi̯] > [roæ] > [ro̯æ] > *roi* [ru̯a] 'king'; Spanish *maíz* 'corn' > dial. *ma̯iz*; *fi.lí.ŏ.lu* > *fi.lio.lu* [fi.li̯o.lo] 'small son' (> Italian *figliuolo*, French *filleul*, Spanish *hijuelo*); (Luick 1891: 338, Jespersen 1897-99 [1904]: § 198, Passy 1891: § 475, Sievers 1893[4]: § 395).[29]

The other element of the diphthong is the glide. It has sometimes been stated that the glide of a diphthong must consist of a high vowel. From the data it is easy to prove that this is not always the case (see also Sievers 1893[4]: § 389; Ruhlen 1975: 52). However, it can be safely stated that the majority of diphthongs and specially the most frequent ones tend to have a high glide, cf. (D6) and section 5.2., which agrees with Stampe's idea that the best glide is maximally colored.

On the other hand, it is commonplace in the phonetic literature that the actual realization of these glides is not as extreme as the transcriptions would suggest (Sievers 1893[4]: § 389; Jespersen 1897-99 [1904]: § 212; Solomon and Sara 1984; Jha 1985). In fact, the realization of the glide presents different degrees of accuracy

[29] Cf. a detailed discussion of the process iə̯ > je in the Romance languages in Sánchez Miret (1996; in press).

in different styles: see, for example, the situation of the pronoun *I* in English below, (Stampe 1972: 585).

very careful speech		ordinary speech		casual speech
[a̯i]	→	[a̯e]	→	[a:] (*I'm here* [a:mir])

Equally the German diphthongs (e.g., *Zeit* 'time', *Haus* 'house', *Scheu* 'fear') are usually transcribed as /ai, au, oi/, /aɪ, aʊ, ɔʏ/ or /aɛ, ɑɔ/, and one can also find the phonetic transcription [ae, ao, ɔø] (Mioni 1973: 149; Iivonen 1989: 24-27; Moosmüller 1997). We see that linguists have used four possible degrees of height for the transcription of the glide.

In conclusion, we see in these facts, first the tendency to the polarization in a sonorous nucleus and a colored glide (whose results are tentatively represented in the current phonological transcriptions of diphthongs), and then the action of postlexical assimilation processes, which are responsible for the observed less extreme realization of glides, and also of nuclei.

5.4. Is there any preference for the position of the nucleus?

Finally, I want to address the question of preferences in the position of the nucleus relative to the glide. Traditionally, diphthongs are considered to be sequences of nucleus + glide, that is, falling diphthongs (see section 2). In fact, this seems to be the most frequent form of a diphthong. The distribution of falling and rising diphthongs in the data is shown in Table 6.[30]

Table 6. The distribution of falling and rising diphthongs.

source	type frequency		token frequency	
	falling	rising	falling	rising
UPSID	19	10	93	25
Weeda	41	24	65	35
SPhA monoph	4	6	4	10
SPhA alloph	26	12	47	14
SPhA bipho	28	6	48	8
SphA total			99	32

The reason for this can be seen again in the process of spontaneous diphthongization. As Donegan (1978 [1985]: 196) states: "The original diphthongization of a simple vowel [...] typically produces a falling diphthong: V(:) → VV̯." This is so because a falling diphthong generally has a greater duration than the equivalent ris-

[30] The sources do not always clearly state the syllabicity of the diphthongs. I count as rising diphthong every one whose second element has greater sonority than the first one, and as falling those diphthongs with the inverse pattern. I do not count diphthongs with elements of the same sonority: [ui, iu, ʊɪ, oe, eo, ɔɛ, iɨ, ɨi].

ing one, thus making it easier to perceive (see measurements in Gottfried et al. 1993: 218-219, about English diphthongs; Salza et al. 1987, about Italian; Borzone de Manrique 1979, about Spanish).

Falling diphthongs are the normal result of a diphthongization process, since this is a foregrounding process and it is expected to preserve or even increase the duration of the strengthened vowels. Consequently, a rising diphthong is a bad solution for diphthongization, since it implies some amount of shortening (Donegan 1978 [1985]: 101). In fact, falling diphthongization seems to be the most common form of diphthongization in the history of Germanic and Romance languages (Sánchez Miret 1996: § 2.4). Furthermore, we arrive at the same conclusion if we look at the cases of allophonic diphthongization registered by the SPhA; the data are shown in Table 7.[31]

Table 7. Allophonic diphthongization in SPhA.

falling diphtongs				rising diphthongs
a > aj	o̯ > o̯j	ɤ > ɤẘ	i > iə̯ (2)	ɔ > wɛ
æ > æə̯	ø̯ > ø̯ə̯	ɪ > ɪj	i > eə̯	e̯ > ee̯ (2)
ɛ > ɛj	e̯ > ei	ɻ > ɪɻ	i > eɛ̯	o̯ > oo̯
ɛ > ej	e̯ > ej (4)	ʊ > ʊw	i > ɪi	o̯ > wo̯
ɛ > ɛʌ̯	e̯ > eə̯	ʊ > uj	u > uw	o̯ > jo̯
ɛ > ɛə (2)	o̯ > oj	ʊ̯ > ʊ̯ẘ	u > uə̯ (2)	e̯ > je
ɔ > ɔʌ̯	o̯ > ou	i > ij (4)	u > ʊu	o̯ > wo
ɔ > ɔə	o̯ > ow (2)	i > iə̯ (2)	u > ʊə̯	ʊ > wɪ
o̯ > o̯ə̯	o̯ > ɵw	i > iə̯	u > ow	ɯ > ẘə
o̯ > o̯ə̯ (2)	o̯ > o̯w	i > ej (2)	y > yə̯	i > ẘɪ
			ɐ > ɐj	

ẘ = approximant back unrounded

Moreover, the most frequent evolution among diphthongs is falling > rising (and not rising > falling), as in French *ē* > *e̯i* > *o̯i* > *oe̯* > *o̯e* > *u̯e* > *u̯a*: e.g., *rege* > *roi* 'king'; or Spanish *o̯i* > *oe̯* > *o̯e* > *u̯e*: e.g., *cŏriu* > *coiro* > *cuero* 'leather'.[32] The reason is that falling diphthongs are easily reducible[33], and the evolution *falling* > *rising* implies a temporal reduction. On the other hand, rising diphthongs have difficulties in increasing the duration of their glides, which would be their only possible way of changing from rising to falling (Sánchez Miret 1996; in press).

[31] I do not include two cases where the SPhA does not make explicit the syllabicity, which otherwise is not deducible: Amuesha e > ɤe; Mandarin Chinese ɤ > oɤ. I have also excluded the cases alluded above (see fn. 16): Chukchi e > a̯e, a > a̯e, u > o̯u, ə > a̯ə, ɪ > e̯ɪ; Nez Percé ɪ > ə̯ɪ, i > ə̯i.

[32] A counterexample would be the evolution i̯e > iə̯, which is reconstructed for the history of several South Italian dialects, but see in Sánchez Miret (1996; in press) the arguments for the more plausible reconstruction: iə̯ > i̯e.

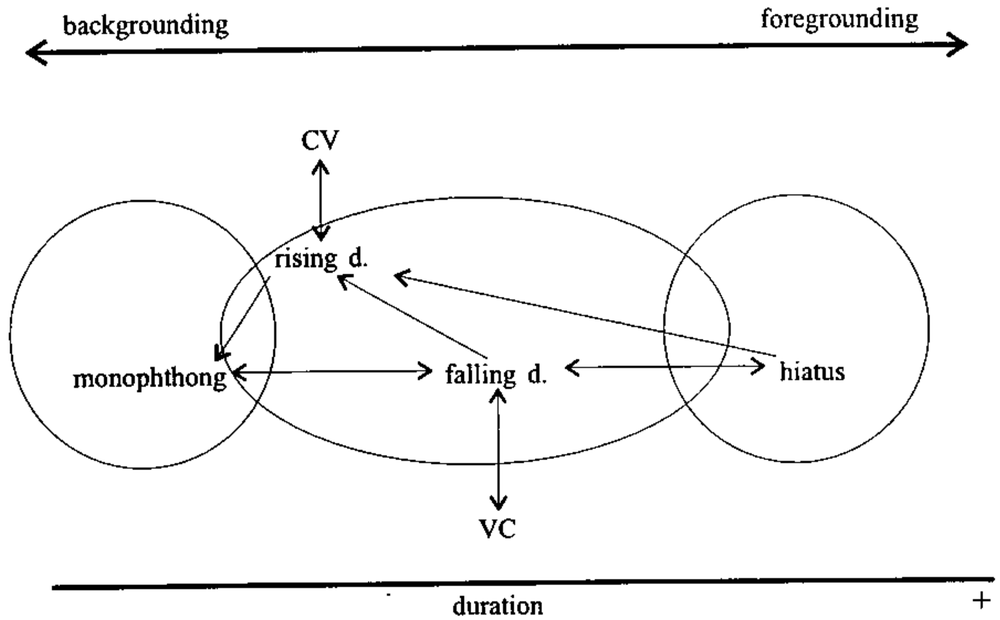
[33] Cf. the results by Gottfried et al. (1993: 219) for American English diphthongs, where the differences in duration between slow vs. fast tempo are of 27% for /aɪ, aʊ/ but only of 19% for /ju/.

The nature of the process of diphthongization points to the conclusion that falling diphthongs are more prototypical and this is confirmed by their frequency in the inventories. However this seems to be in contradiction with what I said about the exigencies of tautosyllabicity (see section 5.1).

6. Conclusion

This contradiction focuses on the conflicting requirements of the category 'diphthong'. I think we can overcome this contradiction if we look at the dynamics of the facts. The diphthongization of a vowel tends to increase its latent duality, and the best way of doing this is for it to develop into a falling diphthong with maximal distant endpoints. Extreme evolutions lead to hiatus or to VC sequences. Furthermore, the realization of a diphthong varies, so that in most casual styles it can be produced as a monophthong or, if the circumstances are given, it can be abbreviated via syllabicity change. The resulting rising diphthong can monophthongize (French *ŏ* > *u̯e* > *œ*; e.g., *cŏr* > *coeur* 'heart') or can evolve into a CV sequence. Some of these evolutions, as such diphthongization and conversion into hiatus are favored in foregrounding contexts (like stressed syllable and phrasal accent). In such contexts also an increase in duration is expected. On the opposite side, falling > rising and diphthong > monophthong are favored in backgrounding contexts, which usually are accompanied by decrease in duration. All these processes take place inside the supracategory of vocalic sounds. The evolution: diphthong > CV, VC implies the apparition of elements of another different supracategory, namely consonantal sounds. This is what I try to represent in Diagram 3.

Diagram 3. Processes inside then supracategory of vocalic sounds.



During these processes some diphthongs can be analyzed as allophonic manifestations of monophthongs, or as phonemic units or as biphonematic sequences. All through these evolutions into the interior or to the exterior of the category the preferences of the diphthong are fulfilled to a different degree. These preferences are represented below.

unity	duality
tautosyllabicity	perceptual distance
cohesion	duration

a) Tautosyllabicity has the aim of ensuring some degree of unity.
b) Syllabic cohesion serves unity.
c) Perceptual distance has the aim of providing some degree of duality. A maximal degree of distance serves perceptibility. A minimal degree serves pronounceability. The process of diphthongization polarizes the incompatible features of a monophthong (color and sonority). The process of assimilation has the opposite function.
d) An optimal diphthong is longer than a short vowel. Falling diphthongs fulfill this condition better than rising ones.
e) The category of diphthong shows a gradation between unity and duality. The prototypical diphthong should exhibit both features to a certain degree.

In conclusion, diphthongs are complex phenomena that show both unity and duality features, and which can arise in many forms in the human languages. I think that we come closer to their understanding if we look at their dynamics.

REFERENCES

Alcina Franch, J. and Blecua, J. M. 1975. *Gramática española*. Barcelona: Ariel.

Andersen, H. 1972. "Diphthongization". *Language* 48. 11-50.

Andersen, H. 1973. "Abductive and deductive change". *Language* 49. 567-593.

Andruski, J. E. and Nearey, T. M. 1992. "On the sufficiency of compound target specification of isolated vowels and vowels in /bVb/ syllables". *Journal of the Acoustical Society of America* 91. 390-410.

Avram, A. 1975. "Perceperea secventelor de două vocale sintetice si problema naturii fonetice a diftongilor". *Fonetică si Dialectologie* 9. 67-75.

Berg, T. 1986. "The monophonematic status of diphthongs revisited". *Phonetica* 43. 198-205.

Bertinetto, P. M. 1988. "Felicity and poverty of Experimental Phonology". *Quaderni del Laboratorio di Linguistica* 2. 85-111.

Bertinetto, P. M., Caboara, M., Gaeta, L. and Agonigi, M. 1994. "Syllabic division and intersegmental cohesion in Italian". In Dressler, W. U. et al. (eds.). 19-33.

Bertinetto, P. M. and Loporcaro, M. (eds.). 1988. *Certamen Phonologicum. Papers from the 1987 Cortona Phonology Meeting*. Torino: Rosenberg & Sellier.

Bladon, A. 1985. "Diphthongs: A case study of dynamic auditory processing". *Speech Communication*. 4. 145-154.

Blumstein, S. E. 1991. "The relation between phonetics and phonology". *Phonetica* 48. 108-119.

Borzone de Manrique, A. M. 1979. "Acoustic analysis of the Spanish diphthongs". *Phonetica* 36. 194-206.

Catford, J. C. 1977. *Fundamental problems in phonetics*. Edinburgh: Edinburgh University Press.

Collier, R., Bell-Berti, F. and Raphael, L. J. 1982. "Some acoustic and physiological observations on diphthongs". *Language and Speech* 25. 305-323.

Corrigan, R., Eckman, F. and Noonan, M. (eds.). 1989. *Linguistic categorization*. Amsterdam: John Benjamins.

Cowan, N., Braine, M. D. S. and Leavitt, L. A. 1985. "The phonological and metaphonological representation of speech: evidence from fluent backward talkers". *Journal of Memory and Language* 24. 679-698.

Crothers, J. 1978. "Typology and universals of vowel systems". In Greenberg, J. H. et al. (eds). 93-152.

Crothers, J. H., Lorentz, J. P., Sherman, D. A. and Vihman, M. M. 1979. *Handbook of phonological data from a sample of the World's languages. A report of the Stanford Phonology Archive*. [Introduction. Vol. 1: Phonetic inventories. Vol. 2: Indexes.] Stanford, Department of Linguistics, Stanford University.

Donegan, P. J. 1978 [1985]. *On the natural phonology of vowels*. New York: Garland. [Ph.D. originally published in *Working Papers in Linguistics* 23. Ohio State University. 1978.]

Dressler, W. U. 1984. "Explaining natural phonology". *Phonology Yearbook* 1. 29-51.

Dressler, W. U. 1989. "Markedness and naturalness in phonology: The case of natural phonology". In Tomić, O. M. (ed). 111-120.

Dressler, W. U. 1996. "Principles of naturalness in phonology and across components". In Hurch, B. and R. A. Rhodes (eds). 1996. 41-51.

Dressler, W. U., Prinzhorn, M. and Rennison, J. R. (eds.). 1994. *Phonologica 1992. Proceedings of the 7th International Phonology Meeting*. Torino: Rosenberg and Sellier.

Edström, B. 1971. "Diphthong systems". Stockholm University. Ms.

Gottfried, M., Miller, J. D. and Meyer, D. J. 1993. "Three approaches to the classification of American English diphthongs". *Journal of Phonetics* 21. 205-229.

Greenberg, J. H., Ferguson, C. A. and Moravcsik, E. A. (eds.). 1978. *Universals of human language. 2: Phonology*. Stanford: Stanford University Press.

Harrington, J. and Cassidy, S. 1994. "Dynamic and target theories of vowel classification: Evidence from monophthongs and diphthongs in Australian English". *Language and Speech* 37. 357-373.

Hurch, B. and Nathan, G. 1996. "Naturalness in phonology". *Sprachtypologie und Universalienforschung* 49. 231-245.

Hurch, B. and Rhodes, R. A. (eds.). 1996. *Natural phonology: The state of the art*. Berlin: Mouton de Gruyter.

Iivonen, A. 1989. *Regional German vowel studies*. Helsinki: University of Helsinki. [Helsingin yliopiston fonetiikan laitoksen monisteita. Mimeographed series of the Department of Phonetics, University of Helsinki 15].

Jespersen, O. 1897-99. [1904]. *Lehrbuch der Phonetik*. Leipzig/Berlin, Teubner. [= *Fonetik. En systematisk fremstilling af læren om sproglyd*. Krbenhavn, 1897-99.].

Jha, S. K. 1985. "Acoustic analysis of the Maithili diphthongs". *Journal of Phonetics* 13. 107-115.

Keating, P. A. 1988. "The window model of coarticulation: Articulatory evidence". *UCLA Working Papers in Phonetics* 69. 3-29.

Krüger, F. 1923. *El dialecto de San Ciprián de Sanabria. Monografía leonesa*. Madrid: Sucesores de Hernando.

Ladefoged, P. 1982. *A course in phonetics*. Los Angeles: University of California.

Lass, R. 1984. "Vowel system universals and typology: Prologue to theory". *Phonology Yearbook* 1. 75-111.

Lehiste, I. and Peterson, G. E. 1961. "Transitions, glides and diphthongs". *Journal of the Acoustical Society of America* 33. 268-277.

Lindblom, B. 1986. "Phonetic universals in vowel systems". In: Ohala, J. J. and J. J. Jaeger (eds). 1986. 13-44.
Luick, K. 1891. "Unechte und steigende Diphthonge". *Beiträge zur Geschichte der deutschen Sprache und Literatur* 16. 336-342.
Luschützky, H. C. 1992. *Zur Phonologie der Affrikaten*. Frankfurt a. M.: Wissenschaftliche Buchhandlung Theo Hector. [Forum Phonetikum, 48].
Maddieson, I. 1981. "UCLA phonological segment inventory database. Data and index". *UCLA Working Papers in Phonetics* 53. 1-242.
Maddieson, I. 1984. *Patterns of sounds*. Cambridge: Cambridge University Press. [With a chapter contributed by Sandra Ferrari Disner].
Maddieson, I. 1991. "Testing the universality of phonological generalizations with a phonetically specified segment database: Results and limitations". *Phonetica* 48. 193-206.
Marotta, G. 1987. "Dittongo e iato in italiano: una difficile discriminazione". *Annali della Scuola Normale Superiore di Pisa. Serie III. Classe di lettere e filosofia* 17. 847-887.
Marotta, G. 1988. "The Italian diphthongs and the autosegmental framework". In Bertinetto, P. M. and M. Loporcaro (eds.). 389-420.
Menzerath, P. 1941. *Der Diphtong. Eine kritische und experimentelle Untersuchung*. Bonn: Dümmler.
Mioni, A. M. 1973. *Fonematica constrastiva. Note ed esercizi. Italiano, francese, spagnolo, tedesco, inglese, russo, portoghese, rumeno, neerlandese, altre lingue*. Bologna: Pàtron.
Moosmüller, S. 1997. "The spread of Viennese monophthongization: a case of dialect levelling?". *http://www.esf.org*.
Nathan, G. S. 1986. "Phonemes as mental categories". In Nikiforidou, V. et al. (eds.). 212-223.
Nathan, G. S. 1989. "Preliminaries to a theory of phonological substance: The substance of sonority". In Corrigan, R. et al. (eds.). 55-67.
Nathan, G. S. 1996. "Steps towards a cognitive phonology". In Hurch, B. and R. A. Rhodes (eds.). 105-120.
Nearey, T. M. 1989. "Static, dynamic, and relational properties in vowel perception". *Journal of the Acoustical Society of America*. 85. 2088-2113.
Neary, T. M. and Assmann, P. F. 1986. "Modeling the role of inherent spectral change in vowel identification". *Journal of the Acoustical Society of America* 80. 1297-1308.
Nikiforidou, V., VanClay, M., Niepokuj, M. and Feder, D. (eds.). 1986. *Proceedings of the twelfth annual meeting of the Berkeley Linguistics Society*. Berkeley: BLS, University of California.
Ohala, J. J. and Jaeger, J. J. (eds.). 1986. *Experimental phonology*. Orlando, Florida: Academic Press.
Passy, P. 1891. *Étude sur les changements phonétiques et leurs caractéres généraux*. Paris: Firmin Didot.
Peeters, W. J. M. 1991. *Diphthong dynamics: A cross-linguistic perceptual analysis of temporal patterns in Dutch, English, and German*. Kampen: Mondiss. [Doctoral dissertation, University of Utrecht. CIP—Gegevens Koninklijke bibliotheek: Den Haag].
Peranteau, P. M., Levi, J. N. and Phares, G. C. (eds.). 1972. *Papers from the eighth annual regional meeting of the Chicago Linguitic Sociey*. Chicago: CLS.
Pike, K. L. 1947a. *Phonemics: A technique for reducing languages to writing*. Ann Arbor: University of Michigan Press.
Pike, K. L. 1947b. "On the phonemic status of English diphthongs". *Language* 23. 151-159.
RAE = Real Academia Española. 1973. *Esbozo de una nueva gramática de la lengua española*. Madrid: Espasa-Calpe.
Rischel, J. 1991. "The relevance of phonetics for phonology: A commentary". *Phonetica* 48. 233-262.
Rosetti, A. (ed.). 1959. *Recherches sur les diphtongues roumaines*. Bucarest: Éditions de l'Académie/ Munksgaard.
Ruhlen, M. 1975. *A guide to the languages of the world*. Stanford: Language Universals Project, Stanford University.
Salza, P. L., Marotta, G. and Ricca, D. 1987. "Duration and format frequencies of Italian bivocalic sequences". *Quaderni del Laboratorio di Linguistica* 1. 29-32.

Sánchez Miret, F. 1996. *La diptongación en las lenguas románicas*. Salamanca.. Ph.D. thesis.
Sánchez Miret, F. In press. "Los diptongos [i̯ę, u̯ę] en la Romania y su relación histórica con [i̯e, u̯e, u̯o]".
Schubiger, M. 1977[2] [1989]. *Introducción a la fonética*. Valladolid: Universidad de Valladolid. [= *Einführung in die Phonetik*. Berlin: de Gruyter. 1977[2], 1970[1]]
Sievers, E. 1893[4]. 1901[5]. *Grundzüge der Phonetik zur Einführung in das Studium der Lautlehre der indogermanischen Sprachen*. Leipzig: Breitkopf and Härtel. [1876[1]]
Slama-Cazacu, T. 1958 [1959]. "The experimental reversed speaking, with special view to diphthongs". In Rosetti, A. (ed.). 123-134.
Solomon, I. and Sara, S. J. 1984. "English diphthongs, [ai, oi, au]". In Van den Broecke, M. P. R. and A. Cohen (eds.). 411-414.
Stampe, D. 1972. "On the natural history of diphthongs". In Peranteau, P. M. et al. (eds.). 578-590.
Strange, W. 1989. "Evolving theories of vowel perception". *Journal of the Acoustical Society of America* 85. 2081-2087.
Tomić, O. M. (ed.). 1989. *Markedness in Synchrony and Diachrony*. Berlin: Mouton de Gruyter.
UPSID. 1992. *UCLA Phonological Segment Inventory Database. PC software package*. Los Angeles, UCLA. [Principal investigator: Ian Maddieson. Programmer: Kristin Precoda]
Van den Broecke, M. P. R. and Cohen, A. (eds.). 1984. *Proceedings of the Tenth International Congress of Phonetic Sciences*. Dordrecht: Foris.
Weeda, D. 1983. "Perceptual and articulatory constraints on diphthongs in Universal Grammar". *Texas Linguistic Forum* 22. 147-162.