

A SHORT OCCITAN TUTORIAL

How to read Occitan?

Due to its intermediary position between French, Italian and Spanish, Occitan shares pronuciations with these three languages.

Stress

The rules to know where the stress lies in a word are simple. If a word ends in a vowel or s (plural), the stress is on the syllable before the last one. If the word ends in a consonant except s, the stress lies on the last syllable. All other cases are explicitly written with an accent on the vowel.

Vowels

- a** or **à** is to be read **[a]** as in (southern) French, Spanish and Italian. About the same sound as in English 'fat'. One should notice that an 'a' at the end of the word is the feminine ending, which pronunciation varies from places to places. The most common pronunciation is [o], as 'ò' below.
- e** or **é** stands for the **[é]** sound: French 'été'
- è** stands for the **[è]** sound: French 'belle', italian 'bella'. About the same sound as in English 'get'.
- i** or **í** stands for the **[i]** sound: French 'nid', Spanish 'mi, hijo'. About the same sound as in English 'grid'
- o** or **ó** is to be read **[u]** as in French 'tout', Spanish 'tu'. About the same sound as in English 'foot'.
- ò** stands for the **[o]** sound, as in Spanish 'todo', French 'gros' A sound between English 'got' and 'all'.
- u** or **ú** is to be read **[ü]** as in French 'tu', German 'Tür'.
- á** appears only a the end of words, where it is sounded **[o]**(**[è]** along the mediterranean coast).

Consonants

- p,b,m,f,t,d,n**, have about the same pronunciation as in English More precisely, **p, m ,f, t, n** are sounded as in French, Spanish or Italian i.e.: **[p],[m],[f],[t],[n]**, but notice that **final n** often vanishes, especially in Lengadocian. **b** and **d** are pronounced exactly the same way as in Spanish.
- v** is pronounced **[b]** as in Spanish in general Occitan, as well as in dialects of Gasconha, Lengadoc, e Auvernha. The other dialects read it **[v]**
- s** is pronounced **[s]**, or **[z]** between two vowels, as is the case in French and English. In such a position, **ss** is to be read **[s]**.
- sh** is pronounced **[sh]**as in English (eg Foish [fwish]=Foix town).
- c** is sounded **[k]** before **a,o** and **u**. Before **e** or **i**, it is pronounced **[s]**, as in French or English. **ç**, is also pronounced **[s]**.
- ch** is pronounced **[tsh]** as in English(eg charrar [tshar'ra]= to gossip).
- g** is pronounced **[g]** before **a,o** and **u**. Before **e** or **i**, it is pronounced **[j]**(or **[dj]**), as in French or English. It is sounded **[sh]** at the end of non-greek words (eg pueg =[püèsh] =hill).
- j** is pronounced **[j]** (or **[dj]**).
- l** is pronounced **[l]**, as in French, Spanish. It may be sounded as the 'l' sound in English 'call' at the end of syllables.



- lh** is pronounced [**ly**] (eg: palha =['palyo]=straw). It corresponds to Italian /gli/, Portuguese /lh/, Castilian and Catalan /ll/.
- nh** is pronounced [**ny**] as in Portuguese(eg: vinha ['binyo]=vine). It corresponds to Italian and French/gn/, Portuguese /nh/, Castilian /ñ and Catalan /ny/.
- r** is pronounced [**r**] trilled as in Spanish or Italian. In Provença, it is often sounded [**R**] as in French. Notice that **rr** must be pronounced as two 'r':[**rr**] as in Spanish. .
- th** is pronounced [**ty**], but this is restricted to Gasconha. (eg: bèth ['bèty]=beautiful).

Diphthongs

- ai** stands for [**ay**] as in English 'ride'.
- au** stands for [**aw**] as in English 'out'.
- èi** stands for [**èy**] as in Portuguese 'brasileiro'
- èu** stands for [**èw**].
- ei** stands for [**ei**].
- eu** stands for [**ew**].
- iu** stands for [**iw**].
- òi** stands for [**oy**] as in English 'boy'.
- òu** stands for [**ow**] as in American English 'so'(slightly more open).
- oi** stands for [**uy**] as in French 'rouille'.
- uu** stands for [**üw**] (only in Gascon).

As an exercise, try to read this: Lo sermon del curat de Cucunhan, by Aquiles Mir

The Occitan Article

As in all romance languages, there are two genders (feminine and masculine), and two numbers (singular and plural) for nouns and adjectives, articles, pronouns, etc.

The definite article

For the Standard Occitan, the definite article forms are

Number	masculine	feminine
singular	lo	la
plural	los	las

These are subject to small dialect variations:

- In the area between Tolosa, Foish e Carcassona, the Masculine form are **le**, **les**. These should not be regarded as French influences, since they are present in documents previous to the French presence in this area. They should rather be seen as the evolution of Latin ILLE > le, whereas the evolution has been ILLUM > lo elsewhere.
- In some pyreneans valleys(Coseran, Comenges and Bigorre), the forms are quite different:

Number	masculine	feminine
singular	eth [et, ety]	era
plural	eths [ets]	eras

They are in good agreement with the Gascon dialect evolution (see {10} in the linguistics page), and are transition forms to Castillan (el, la).

In the northern areas, la, las may be sounded [lo, los], along with the general trend to turn non-stressed [a] into [o] in this area (see {21} in the linguistics page).

The indefinite article

The Standard Occitan indefinite article forms are:

Number	masculine	feminine
singular	un	una
plural	de(unes)	de (unas)

The **de** plural form is the true indefinite form; **unes**, **unas** rather mean *some, a few*. For exemple

Fau de còcas means *I am cooking cakes* whether

Fau unas còcas means *I am cooking a few cakes*.

One must mention that Gascon has slightly different forms

Number	masculine	feminine
singular	un	ua [üo, yo]
plural	(no article, as in English)	(no article, as in English)

The demonstratives

In Occitan, there are three degrees of demonstratives (vs two in English -here/there). The degree chosen depends on the distance of the object, and are approximately

- aicì** close to me (here, in my hand)
- aquì** close to you (there, a few steps away, in your hand)
- ailà** close to him (over there, far away, in his/her hand, and he/she can't hear us).

Accordingly, the Standard Occitan demonstrative adjectives and pronouns are the following:

distance	pronoun	number	masculine	feminine
aicì	aiçò	sing	aiceste	aicesta
		plur	aicestes	aicestas
aquì	aquò	sing	aqueste	aquesta
		plur	aquestes	aquestas
ailà	ailò	sing	aquel	aquela
		plur	aqueles	aquelas

there is also a neutral (in terms of distance) pronoun, **çò**, which has many uses in idiomatics.

- In southern dialects, *aicì*, *aicò*, *aiceste*,... are sounded [a'si, a'so, a'seste,...]. There, the masculine plural forms end with (unstressed) [is]. Hence *aicestis*, *aquestis*, *aquelis*.
 - The Tolosan dialect has *aiciu* and *aquiu* instead of *aicì*, *aquì*.
 - In northern areas (close to Lemosin), the **a**-sound at the beginning of these words may vanish. Hence *'quì*, *'quò*, *'quel*, etc.
-

Occitan Verb Conjugation

As in all romance languages, there are three basic conjugations in Occitan.

The verbs of the first conjugation have **-aras** infinitive ending

The verbs of the second conjugation have **-ir** as infinitive ending

The verbs of the third conjugation have **-er** or **-e** as infinitive ending

The different simple tenses are:

Present, preterit, imperfect, futur, imperatiu, subjunctiu present, subjunctiu passat, condicional, particip present, particip passat.

Then, a wide number of composite tenses can be formed using those ones. (The conjugation is the one of the 'Standard Occitan', but there may be interferences with my Tolosenc dialect).

tense	cantar	florir	batre
Present	canti (north and east: cante) cantas canta cantam cantatz cantan	florissi (north and east: florisse) florisses florìs florissèm florissètz florissen (or -isson)	bati (north and east: bate) bates bat batèm batètz baten (or baton)
Imperfait	cantava (or -avi) cantavas cantava cantavem cantavetz cantavan	florissiá florissiás florissiá florissiam florissiatz florissián	batiá batiá batiá battiam batiatz batián
Futur	cantarà cantaràs cantarà cantarem cantaretz cantaràn	florirà floriràs florirà florirem floriretz floriràn	batrà batràs batrà batrem batretz batràn
Preterit	cantèri (south: cantègui) cantères (south: cantègues) cantèt (south: cantèc) cantèrem (south: cantèguem) cantèretz (south: cantèguetz) cantèron (south: cantègon)	floriguèri (south: florigui) floriguères (south: florigues) floriguèt (south: floric) floriguèrem (south: floriguem) floriguèretz (south: floriguetz) floriguèron (south: florigon)	batèri (south: batègui) batères (south: batègues) batèt (south: batèc) batèrem (south: batèguem) batèretz (south: batèguetz) batèron (south: batègon)
Subjonctiu present	cante cantes cante cantèm cantètz canten	florisca floriscas florisca floriscam floriscatz floriscan	bata batas bata batam batatz batan
Subjonctiu present	cantèssi cantèsses cantèsse (north: cantès) cantèssèm cantèssetz cantèssen (or -èsson)	floriguèssi floriguèsses floriguèsse (north: floriguès) floriguèssèm floriguèssetz floriguèssen (or -èsson)	batèssi batèsses batèsse (north: batès) batèssèm batèssetz batèssen (or -èsson)
condicional	cantariá cantariás cantariá cantariam cantariatz cantarián	floririá floririás floririá floririam floririatz floririám	batriá batriás batriá batriam batriatz batrián

Note:

□ "south" means *under pyrenean influences*, i.e. mostly in dialects of Tolosa, Foish and further west and south. It must be added that in these areas, the -iss- combination in all forms of 2nd group verbs,

like *florir*, may be uttered [-ish] or even [-eysh].

- "east" means *under provençal influence*, i.e. in dialects of Montpelhièr and further east.
- "North" means *under lemosin or Auvenhat influence*, i.e. in dialects s of Orhac, Sarlat, etc...

Back to [Occitan Language Page](#)

OCCITAN LINGUISTICS

PHONETICS AND GRAPHIC SYSTEM

□ Vowels

there are eight vowels in Occitan, which can be stressed or unstressed.

-[a] is written 'a'

-[e] is written 'e'

-[i] is written 'i'

-[u] is written 'o'

-[ü] is written 'u'

-[è] is written 'è'. This sound exists only as stressed vowel. If for some reason, the stress changes places, it becomes [e], written 'e'.ex:

'apèla' (he calls) becomes 'apelar' (to call), when adding '-ar' which bears the stress, to form the infinitive.

-[o] is written 'ò'. This sound exists only as a stressed vowel. If for some reason, the stress changes places, it becomes [u], written 'o'.ex:

'pòrta' (he carries) becomes 'portar' (to carry), when adding '-ar' which bears the stress, to form the infinitive.

-there is also an atonic vowel, which is essentially the feminine ending. This vowel varies.

The most common pronunciation is [o], but some areas in the pyrenees, and near Montpelhièr pronounce it [a], as in the middle ages, and along the Atlantic coast, it is sounded [ö]. This sound is written 'a', as it used to be in the middle ages, and as it is still the case in Catalan.

□ Diphtongs

Occitan has a great number of diphtongs. It is a main difference with the other romance language, Occitan is one of the most conservative language on this point.

-[ay] is written 'ai' ex: gaitar: to look

-[aw] is written 'au' ex: aut,naut: high

-[ey] is written 'ei' ex: veire: to see

-[ew] is written 'eu' ex: beure: to drink

-[èy] is written 'èi' ex: lèit: bed

-[èw] is written 'èu' ex: lèu: quick

-[iw] is written 'iu' ex: riu: river

-[òw] is written 'òu' ex: nòu: new

-[òy] is written 'òi' ex: còire: to cook

-[uy] is written 'oi' ex: coire: leather

-[üw] is written 'uu' ex: cuu: bottom(Gascon)

□ Glides

There are three glides:

-[y], written i: ieu=[yew,yu] : I, me

-[w], written o: bois=[bwis, bwish]:box (plant)

-[ÿ], as in french 'lui', is written u: nuèit=[nÿèt]: Night. One can notice that this glide tends too disappear in many dialects, or to be replaced by [y]:

buðu=[bÿðu]>[byðu],nuèit=[nÿèt]>[nèt].

□ Consonnants.

We give here the richer consonantic system to be found in Lengadocian dialects. We then describe how it can be reduced according to various dialects.

labials

[p], written 'p'.

[b] written 'b', if it comes from latin P, written 'v' if it comes from latin B or V. Note that when it is written 'v', Lemosin and Provençal pronounce it [v].

[m] is written 'm'

labio-dentals

[f] is written 'f', even in greek words.

[v] see before.

dentals

[t] is written 't'

[d] is written 'd'

[n] is written 'n'

[s] is written 's','ss' between two vowels if coming from latin S or SS. It is written 'c' before e and i, and 'ç' before a,o,u, or at the end of words, if coming from a latin C.

[z] is written 'z','s' between two vowels.

palatals

[sh] is written 'sh' in gascon, '-iss' in Lengadocian. When coming from a latin word ending by -DIU (like PODIUM), it is written -g (PODIUM>pueg=[pÿesh]: hill

[j] (as English meaSure) is written 'j' before a,o,u and 'g' before e and i.

[tsh] is written 'ch'

[dj] is written 'tj' before a,o,u and 'tg' before e and i.

[ny] (as French -gn-, Spanish ñ) is written 'nh'

[ly] (as Spanish -ll-) is written 'lh'

velars

[k] is written 'c'before a,o,u and 'qu' before e and i.

[g] is written 'g'before a,o,u and 'gu' before e and i.

[kw] is written 'qu' (exists only before a, only in Gascon and dialects close to Catalan).

[gw] is written 'gu' (exists only before a, only in Gascon and dialects close to Catalan).

liquids

[l] is written 'l'

[r] is written 'r'. This sound is the same as in spanish or italian (trilled r, with the tip of the tongue). Under french influence, it tends to be pronounced as an alveolar r, as in french or german. This phenomenon is strongest in the east and in the younger generations.

[rr] is written 'rr'.

[ll] is written either 'll' or 'tl' according to writer's taste. ('tl' is the same as in catalan, and makes it easier to come from a language to the other)

- A few transcribed texts

PHONETIC EVOLUTION

The evolution from latin

can be easily discribed:

-Classical latin to late spoken latin (almost common to all romance languages evolution):

-Simplification of the declension system to 2 cases

-Disappearance of deponent conjugation

-evolution of the opposition between short/long vowels to the opposition between open/shut vowels

a long, a short -> a

e long -> closed e (french é)

e short -> open è (french è)

i long -> i

i short -> closed e (french é)

o long -> closed o (as in 'saw', french 'beau')

o short -> open o (as in french 'pomme')

u long -> u (as in 'foot', french ou)

u short -> closed o (as in 'saw', french 'beau')

-C+E and C+I pronounced [tse],[tsi]

-G+E and G+I pronounced [dje],[dji]

hence,

-Spoken latin to occitan

-No more declension

-Vanishing of final vowel except a.

-the u sound turns to [ü] like in french or german

-later, the shut o sound turns to u sound, so the present vocal system is:

ü

a o u

è é i

The specificities of occitan

as opposed to other romance languages are:

INNER EVOLUTIONS, DIALECTS

Little is known about the different dialects in the middle ages. They must have existed, since several differences between the dialects spoken nowadays seem to have deep roots in the past, but almost all texts of the medieval period are written in some sort of common-tongue, used as well by poets ("Trobadors") as by administrators and lawyers so the knowledge about the spoken tongue is rather tiny.

There are now six major dialects, which are:

Southern dialects

- Lengadocian (fr:Langudocien) [Ethnologue Database](#)
- Gascon [Ethnologue Database](#)
- provençal [Ethnologue Database](#)

Northern dialects

- Lemosin (fr:limousin) [Ethnologue Database](#)
- Auvernhat (fr:Auvergnat) [Ethnologue Database](#)
- Vivaro-Alpin

The differences between these dialects can be made clear by considering a three-poles-hypothesis, from where the innovations would have scattered:

- the Pyenees area
- the flat northern half of France

-Massis Septentrional (fr: Massif Central), but the role of this area may have only been to stop the evolutions coming from the north, since it is a mountainous area, where communications are not easy.(see below CT>CH)

These evolutions should be added to the description of the evolution from latin described in the previous section.

□ THE PYRENEAN POLE

This is the area where dialects have been most influenced by the formerly spoken non-indoeuropean languages. As a matter of fact, the gaul population in this area had been, at most, very recent by the time of the roman conquest, and in many places, the previous languages, where still in use. (Julius Caesar wrote: GALLOS AB AQUITANIS GARUMNA FLUMEN DIVIDIT) They share certain phonetical evolutions with Euskara (Basque) and Iberic dialects like Aragones and Castellano (standard spanish). Actually, these languages may have been those of the Basques' ancestors. (For information about more recent influences between [Basque and Occitan](#) -Gascon dialect.) The influences coming from this pole are:

- a general strong conservatism of the latin consonantic system
- the sound [i] as the mark of the first person of the singular{1}:
canti: I am singing, cantavi: I was singing, cantèri: I sang

-absence or degradation of labio-dental consonants f and v:
v is always a voiced bilabial sound, plosive or approximant
f has been kept in the eastern part, while it has turned to h sound (like in english) in the western part.

VINU -> vin pronounced [bi],[bin] (everywhere){2},
HABEBAT>aveva [a'wewo] (west), [a'bebo](east) {3}
FILIA -> "hilha"(west), or "filha"(east) pronounced ['hiljo],[filjo].{4}

-the vanishing of intervocalic N{5}:

UNA,LUNA,SEMENARE>ua,lua,semear(semiar)['üo, 'lïo,seme'a(sem'ya)]
This phenomenon also occurred in Portuguese, and maybe in Basque too. It may be related to the vanishing of all final n in Lengadocian and Catalan.

-the latin group -CT- has turned to -IT(-){6}.

LECTU>lèit, DICTU>dit OCTO>ueit

-the AI diphthong tends towards [ei]: LACTE -> lait [lèit]{7}

-the latin groups -IS-, -SC- and -CS- turn to -SH- sound:{8}

PISCE> peish, LAXARE>laishar, AXE>aish, *PISSARE->pisshar

-the latin group -LL- has two different evolutions. In the east, it goes to [lj] sound, like in Catalan. Thus

BELLA,CASTELLU>bèlha, castèlh pronounced [bèljo, kastèlj].{9}

In the west BELLU>*BELL>beth [bètj], whereas BELLA>bèra [bèro]. This is to be compared to the evolution from Aquitanian ILLI (town) to the Basque IRI (town, village).{10}

-The use of the preposition 'a' to introduce a 'thinking object complement'(as in spanish), e.g

Vesi un ostal: I can see a house

Vesi a un òme: I can see a man. {11}

-the last point is the large use of words like "que","ja","be","e" to give more strength to the verb in the sentences as this exists with "que" and "ya" in Spanish. This is most developed in the south western part, where "que" is almost necessary to use any verb. Que veni=I'm coming(litt. That I' coming).{12}

-One should also notice that the evolution of latin U in [ü] is not general or complete in areas under pyrenean influence:

.some dialects [Narbona (Narbonne) Savès:Hòrgas (Forgues) Bragairac, Lombès] pronounce it [ö]{13}

.the western dialects have no [üe] diphthong, they have a [we] diphthong instead.{14}

This pyrenean influences are very strong within the following limits: the pyrenees range in the south, except at the westmost part, where is the Pays Basque, and the eastern part where is the Northern-Catalan area. In the north, the limit is a line from Bordèu to Narbona (Bordeaux-Narbonne). The eastern limit of the F>H, LL>TH/R,AVA>awa evolutions starts by the city of Tolosa (Toulouse), and goes southward approximately in a straight line to the spanish border. Of course, all these phenomena do not stop at the same place, and there are interference bands. The western dialects are traditionally called "Gascon" dialects. The eastern dialects are the southern part of the traditionally called "Lengadocian pyreneenc(Languedocien Pyrénéen)" dialects.

But some spread further away. To have more details about this see the [map](#)

of pyrenean influences.

□ **THE NORTHERN POLE**

The innovations assumed to come from the flat northern part of France are often shared with (old) french. They are stronger in the north, but some have spread as far as the mediterranean coast: see the [map](#) of their extention. They are:

-the evolution since the VIth century of latin CA,GA towards CHA,JA pronounced [tsha],[dja]. This is shared with old french. Thus{ 15}
CANTARE>chantar, CASTELLU>chastèl, GALLINA>jalina.

-the sound [e] as the mark of the first person of the singular{ 16}:
cante: I am singing, cantave: I was singing, cantère: I sang

-the vocalization of all L not followed by a vowel:{ 17}
CASTELLU>chastèl> chastèu, SALTARE> sautar

-the evolution of s-sound followed by p,t or k into a sound [x] lying between [h] and [y] (greek khi), and then its vanishing{ 18}:
ESCOLA>[exkola]>ecola, espiar>epiar, chastèu>chatèu
This is shared with old french.

-the vanishing of intervocalic latin D, instead of its evolution towards a [z] sound in the southern dialects:{ 19}
SUDARE,NUDA>suar,nua instead of susar, nusa in the south.
this is shared with (old) french

-the disappearance of all final consonants{20}
bèc> [bè], còp>[ko], nuech>[nüè],[nyö]

-the trend to turn the unstressed [a] sound into an [o] sound{21}
hence caval >[ko'val],[tsho'val]

As most of these specificities are common with old french, Northern occitan can really be regarded as a transition language between the southern occitan system and the french system, making the change rather continuous, as it should be expected in an area where no major language change has occurred since 20 centuries.

The limit between northern occitan and french is taken conventionnally to be the line where long free latin A becomes [e] (french side) or remains [a] (occitan side). The limit between northern occitan and southern occitan is the line where CA becomes CHA.

□ **THE Massis-Septentrional POLE**

There are two phenomena that can be thought to be specific to the Massis-Septentrional zone, and that may be related with one another:

-the evolution of the latin group CT (and some times PT)into -ch{22}:
LACTE>lach, NOCTE> nuoch, nuech, LECTU>lièch, FACTA>facha.

Only spanish and some northern italian dialect present the same evolution. This may be due to some old population groups remaining in these rocky areas. It may be related to the palatalization trend described in the northern occitan section above. The limits of the phenomenon are the Bordèu-Narbona line in the south and a bow cutting in the middle the northern occitan area in the north.

-the trend to palatalyse all consonants{23}:

s>sh , d>dy or dj, t>ty or tsh, v>vy, n>ny, etc...

ex: vinha >['vyinya], libre ['lyibre], tuba > ['tshüba], sonar [shu'na] etc

This evolution has spread on a much smaller territory, essentially Auvernha.

□ THE FEATURES OF THE MAIN DIALECTS

-Gascon has the features {1},{2},{3},{4},{5},{6},{7},{8},{10},{11},{12},{13} in Savès, {14} and {17}

-Southern Lengadocian has the features{1},{2},{6},{8},{9},+{7}in the Pyrenees Mountains, +{12} in the area of Narbona

-Northern Lengadocian has the features{1},{2},{22},+{21} in the north, +{17} in the east.

-Provençal has the features {16},{17},{20},{22}, +{19} around Niça.

-Limousin has the features {6} in the north {23} in the south,{15},{16},{17},{18},{19},{20},{21}

-Auvergnat has the features {23},{15},{16},{2} in the south,{17},{18},{19},{20},{21},{23}

-Vivro-Alpin has the features

references

Alibèrt (Loïs) Grammatica Occitana (segon los parlars lengadocians), Institut d'Estudis Occitans, Montpellier 1976(Toulouse 1935)

Bèc (Pèire)

La langue occitane, Que Sais-je? #1059, Presses Universitaires de France, Paris 1963, 1995

Manuel pratique de philologie romane,Presses Universitaires de France, Paris 1971

Darrigand (Robert) Initiation au Gascon, Per Nòste, Pau 1974

Seguy (Jean) Atlas Linguistique et ethnographique de la Gascogne, Editions du CNRS, 1954-1964.

Ravier (Xavier) Atlas Linguistique du Languedoc Occidental,Editions du CNRS, 1984.

Back to [Occitan Language Page](#)