

Format TEI des lexiques du Congrès

Version 10

Aure Séguier, développeuse au Congrès

Octobre 2023



Table des matières

1 Balises.....	1
1.1 Liste des balises, des attributs et des valeurs (version explicative).....	1
1.2 DTD (version technique).....	7
2 Structure générale des entrées.....	13
3 Structure détaillée des entrées.....	14
3.1 L'entrée.....	14
3.2 Le mot vedette.....	14
3.2.1 La nature.....	18
3.2.2 L'orthographe.....	19
3.2.3 Les informations bibliographiques.....	20
3.2.4 Les collocations.....	24
3.2.5 Les formes fléchies et élidées.....	26
3.2.6 Les informations grammaticales.....	29
3.2.6.1 Les catégories grammaticales normalisées.....	30
3.2.6.2 L'affichage des catégories grammaticales.....	31
3.2.6.3 La balise <iType>.....	31
3.2.7 La ou les prononciations.....	33
3.2.8 La syllabisation.....	35
3.2.9 Les informations d'usage.....	35
3.2.10 Les informations de localisation.....	38
3.2.11 Informations diverses.....	39
3.3 La personne-vedette.....	41
3.4 Les informations étymologiques.....	46
3.5 Les différentes acceptions.....	48
3.5.1 Les informations d'usage.....	50
3.5.2 La définition.....	50
3.5.3 Les traductions et références.....	50
3.5.4 Les renvois.....	51
3.5.5 Les sous-entrées.....	56
3.6 Indiquer un fichier lié.....	58
3.7 La balise <note>.....	59
4 Les balises <reg>, <orig> et <expan> et les attributs « norm », « orig » et « expand ».....	60
5 Normes pour les catégories grammaticales.....	62
5.1 Jeu d'étiquettes du Congrès.....	62
5.1.1 Verbes.....	62
5.1.2 Noms.....	65
5.1.2.1 Noms communs.....	65
5.1.2.2 Noms propres.....	65
5.1.3 Adjectifs qualificatifs.....	65
5.1.4 Pronoms, déterminants et adjectifs autres que qualificatifs.....	66
5.1.4.1 Définis.....	66
5.1.4.2 Indéfinis.....	66
5.1.4.3 Démonstratifs.....	68
5.1.4.4 Exclamatifs.....	69
5.1.4.5 Interrogatifs.....	69
5.1.4.6 Possessifs.....	70

5.1.4.7 Relatifs.....	73
5.1.4.8 Personnels.....	73
5.1.4.9 Numéraux cardinaux.....	79
5.1.4.10 Numéraux ordinaux.....	80
5.1.4.11 Partitifs.....	81
5.1.4.12 Cas des articles contractés.....	81
5.1.5 Adverbes.....	83
5.1.6 Conjonctions.....	84
5.1.7 Prépositions.....	84
5.1.8 Interjections.....	85
5.1.9 Particules.....	85
5.1.10 URL.....	85
5.1.11 Ponctuations.....	86
5.1.12 Catégorie grammaticale inconnue.....	86
5.1.13 Catégories de famille grammaticale.....	86
5.2 Utilisation de la norme eagles.....	91
5.2.1 Noms.....	91
5.2.2 Verbes.....	92
5.2.3 Adjectifs.....	93
5.2.4 Pronoms et déterminants.....	94
5.2.5 Articles.....	96
5.2.6 Adverbes.....	96
5.2.7 Adpositions.....	97
5.2.8 Conjonctions.....	97
5.2.9 Numéraux.....	97
5.2.10 Interjections.....	98
5.2.11 Unique.....	98
5.2.12 Résiduel.....	99
5.2.13 Ponctuations.....	99
5.3 Utilisation de la norme Grace.....	99
6 Liste des valeurs possibles pour les attributs « norm ».....	100
6.1 Pour la balise <case>.....	100
6.2 Pour la balise <gen>.....	100
6.3 Pour la balise <gram type="polarity">.....	100
6.4 Pour la balise <iType type="accord">.....	100
6.5 Pour la balise <iType type="altvb">.....	100
6.6 Pour la balise <iType type="inchoative">.....	101
6.7 Pour la balise <lbl type="nature"> (sous <form>).....	101
6.8 Pour la balise <lbl type="reference"> (sous <xr>).....	101
6.9 Pour la balise <mood>.....	102
6.10 Pour la balise <number>.....	102
6.11 Pour la balise <per>.....	102
6.12 Pour la balise <pos>.....	102
6.13 Pour la balise <subc>.....	103
6.14 Pour la balise <tns>.....	103
6.15 Pour la balise <usg type="dom">.....	103
6.16 Pour la balise <usg type="plev">.....	105
6.17 Pour la balise <usg type="register">.....	105
6.18 Pour la balise <usg type="style">.....	106

6.19 Pour la balise <usg type="time">.....	106
6.20 Pour les indications de catégories grammaticales.....	106
6.20.1 Contenu des labels des catégories détaillées.....	106
6.20.2 Contenu des labels des catégories générales.....	107
6.20.3 Etiquettes des différents pronoms personnels.....	107
7 Annexes.....	119
7.1 Codes langue de l'IANA (balises <usg type="lang">, attributs xml:lang).....	119
7.2 Comment choisir entre unité grammaticale et phrase.....	120
7.2.1 De manière générale.....	121
7.2.2 Adverbes.....	121
7.2.3 Adjectifs.....	124
7.2.4 Noms propres.....	125
7.3 Comment étiqueter le type des sous-entrées <re> des dictionnaires ?.....	125
7.3.1 Attestation (attestation).....	125
7.3.2 Citacion (citation).....	125
7.3.3 Collocation (collocation).....	125
7.3.4 Dérivé (derived).....	126
7.3.5 Exemple (example).....	126
7.3.6 Expression (expression).....	126
7.3.7 Gentilé (demonym).....	126
7.3.8 Interjection (interjection).....	127
7.3.9 Locution (locution).....	127
7.3.10 Mot composé (compound).....	127
7.3.11 Nom propre (proper).....	127
7.3.12 Onomatopée (onomatopeia).....	127
7.3.13 Proverbe (proverb).....	127
7.3.14 Verbe pronominal (pronominal).....	127
8 Modèle de document.....	128

1 Balises

1.1 Liste des balises, des attributs et des valeurs (version explicative)

0/1/+ : Aucun, un ou plusieurs

Balise	Contient	Est contenu dans	Attributs
<code><addName></code>	• du texte brut	• <code><persName></code>	• sort (valeurs : nombre entier) ; obligatoire • type (valeurs : honorific, epithet, nick, stagename) ; obligatoire
<code><address></code>	• 1 <code><postCode></code>	• <code><location></code>	
<code><author></code>	• 1 <code><choice></code>	• <code><bibl></code>	• key (identifiant international d'auteur) ; facultatif
<code><bibl></code>	• 0/1/+ <code><author></code> • 0/1/+ <code><biblScope></code> • 0/1/+ <code><country></code> • 0/1/+ <code><date></code> • 0/1/+ <code><edition></code> • 0/1/+ <code><editor></code> • 0/1/+ <code><idno></code> • 0/1/+ <code><note></code> • 0/1/+ <code><ptr></code> • 0/1/+ <code><publisher></code> • 0/1/+ <code><pubPlace></code> • 0/1/+ <code><region></code> • 0/1/+ <code><title></code>	• <code><etym></code> • <code><form></code> • <code><xr></code>	• xml:lang (valeurs : codes de langue et subtags de variantes de l'IANA, voir en 7.1) ; facultatif
<code><biblScope></code>	• 1 <code><choice></code>	• <code><bibl></code>	• unit (valeurs : chapter, column, issue, line, page, part, verse, volum) ; obligatoire
<code><birth></code>	• 0/1 <code><date></code> • 0/1 <code><location></code>	• <code><person></code>	
<code><case></code>	• rien	• <code><gramGrp></code>	• norm (valeurs : voir en 6.1) ; obligatoire
<code><choice></code>	• 1 <code><reg></code> • 0/1 <code><orig></code> • 0/1 <code><expan></code>	• <code><author></code> • <code><biblScope></code> • <code><country></code> • <code><date></code> • <code><edition></code> • <code><editor></code> • <code><gloss></code> • <code><mentioned></code> • <code><pubPlace></code> • <code><publisher></code> • <code><region></code> • <code><settlement></code> • <code><title></code>	

Balise	Contient	Est contenu dans	Attributs
<code><cit></code>	<ul style="list-style-type: none"> 0/1/+ <code><etym></code> 0/1/+ <code><form></code> 0/1/+ <code><note></code> 0/1/+ <code><person></code> 0/1/+ <code><usg></code> 	<ul style="list-style-type: none"> <code><sense></code> <code><xr></code> 	<ul style="list-style-type: none"> type (valeurs : <code>reference</code>, <code>translation</code>) ; obligatoire
<code><colloc></code>	<ul style="list-style-type: none"> du texte brut éventuel 	<ul style="list-style-type: none"> <code><form></code> 	<ul style="list-style-type: none"> norm (valeurs diverses) ; obligatoire orig (texte brut) ; facultatif expand (texte brut) ; facultatif type (valeurs : <code>pre</code>, <code>post</code>) ; obligatoire subtype (catégorie détaillée) ; facultatif value (texte brut) ; obligatoire
<code><country></code>	<ul style="list-style-type: none"> 1 <code><choice></code> 	<ul style="list-style-type: none"> <code><bibl></code> <code><location></code> 	<ul style="list-style-type: none"> key (valeurs : codes ISO de pays) ; obligatoire
<code><date></code>	<ul style="list-style-type: none"> 1 <code><choice></code> 	<ul style="list-style-type: none"> <code><bibl></code> <code><birth></code> <code><death></code> 	<ul style="list-style-type: none"> when (valeurs : date au format normalisé) ; obligatoire
<code><death></code>	<ul style="list-style-type: none"> 0/1 <code><date></code> 0/1 <code><location></code> 	<ul style="list-style-type: none"> <code><person></code> 	
<code><def></code>	<ul style="list-style-type: none"> du texte brut éventuel 	<ul style="list-style-type: none"> <code><sense></code> 	<ul style="list-style-type: none"> xml:id (valeurs : clé unique) ; obligatoire norm (valeurs diverses) ; obligatoire orig (texte brut) ; facultatif expand (texte brut) ; facultatif xml:lang (valeurs : codes de langue et subtags de variantes de l'IANA, voir en 7.1) ; obligatoire
<code><edition></code>	<ul style="list-style-type: none"> 1 <code><choice></code> 	<ul style="list-style-type: none"> <code><bibl></code> 	
<code><editor></code>	<ul style="list-style-type: none"> 1 <code><choice></code> 	<ul style="list-style-type: none"> <code><bibl></code> 	<ul style="list-style-type: none"> key (autorité) ; facultatif role (valeurs : <code>compiler</code>, <code>illustrator</code>, <code>translator</code>...) ; obligatoire
<code><entry></code>	<ul style="list-style-type: none"> 0/1/+ <code><etym></code> 0/1/+ <code><form></code> 0/1/+ <code><note></code> 0/1/+ <code><person></code> 0/1/+ <code><ptr></code> 0/1/+ <code><re></code> 0/1/+ <code><sense></code> 0/1/+ <code><xr></code> 	<ul style="list-style-type: none"> rien 	<ul style="list-style-type: none"> xml:id (valeurs : clé unique) ; obligatoire cert (valeurs : <code>confirmed</code>, <code>current</code>, <code>deleted</code>, <code>not_ok</code>, <code>ok</code>) ; obligatoire
<code><etym></code>	<ul style="list-style-type: none"> 0/1/+ <code><bibl></code> 0/1/+ <code><gloss></code> 0/1/+ <code><lang></code> 0/1/+ <code><mentioned></code> 0/1/+ <code><note></code> 0/1/+ <code><usg></code> 	<ul style="list-style-type: none"> <code><cit></code> <code><entry></code> <code><re></code> 	<ul style="list-style-type: none"> corresp (xml:id d'une forme ou d'un étymon précédé de #) ; obligatoire
<code><expan></code>	<ul style="list-style-type: none"> du texte brut 	<ul style="list-style-type: none"> <code><choice></code> 	

Balise	Contient	Est contenu dans	Attributs
<code><forename></code>	• du texte brut	• <code><persName></code>	• sort (valeurs : nombre entier) ; obligatoire
<code><form></code>	• 0/1/+ <code><bibl></code> • 0/1/+ <code><colloc></code> • 0/1/+ <code><form></code> • 0/1/+ <code><gramGrp></code> • 0/1/+ <code><lbl></code> • 0/1/+ <code><location></code> • 0/1/+ <code><note></code> • 1 <code><orth></code> • 0/1/+ <code><pron></code> • 0/1/+ <code><ptr></code> • 0/1/+ <code><usg></code> • 0/1/+ <code><syll></code>	• <code><cit></code> • <code><entry></code> • <code><form></code> • <code><re></code>	• cert (valeurs : current, not_ok, not_ok_var, ok) ; obligatoire • n (identifiant homographe) ; obligatoire • subtype (valeurs : morpheme, phrase, symbol, word) ; obligatoire • type (valeurs : elided, inflected, inflectelided, lemma) ; obligatoire • value (texte brut) ; obligatoire • xml:id (valeurs : clé unique) ; obligatoire • xml:lang (code langue de l'IANA) ; obligatoire
<code><gen></code>	• rien	• <code><gramGrp></code>	• norm (valeurs : feminine, masculine, neuter) ; obligatoire
<code><gender></code>	• 1 <code><choice></code>	• <code><person></code>	• value (valeurs : feminine, masculine, other) ; obligatoire
<code><genName></code>	• du texte brut	• <code><persName></code>	• sort (valeurs : nombre entier) ; obligatoire
<code><geo></code>	• du texte brut	• <code><location></code>	
<code><gloss></code>	• 1 <code><choice></code>	• <code><etym></code>	• corresp (xml:id d'un <code><mentioned></code> précédé de #) ; obligatoire s'il y a des <code><mentioned></code> • type (valeurs : definition, transcription, translation, type, divers) ; obligatoire
<code><gram></code>	• du texte brut éventuel (sauf pour le type polarity)	• <code><gramGrp></code>	• norm (catégorie grammaticale normalisée, affirmative/negative) ; obligatoire • type (valeurs : catGen, catDet, catFam, catEagles, catGrace, polarity) ; obligatoire
<code><gramGrp></code>	• 0/1 <code><case></code> • 0/1 <code><gen></code> • 0/1-6 <code><gram></code> • 0/1/+ <code><iType></code> • 0/1 <code><mood></code> • 0/1/+ <code><note></code> • 0/1 <code><number></code> • 0/1 <code><per></code> • 0/1 <code><pos></code> • 0/1 <code><subc></code> • 0/1 <code><tns></code>	• <code><form></code>	
<code><hi></code>	• du texte brut	• <code><note></code>	• style (valeurs : style au format css) ; obligatoire

Balise	Contient	Est contenu dans	Attributs
<code><idno></code>	• du texte brut	• <code><bibl></code> • <code><person></code>	• type (valeurs : ISBN, ISSN, DOI, QID, URI, VIAF, ESTC, OCLC pour les bibliographies ; Wikidata, etc. pour les personnes) ; obligatoire
<code><iType></code>	• du texte brut éventuel	• <code><gramGrp></code>	• norm (valeurs : voir en 6.3, 6.4, 6.5 et 6.6) ; obligatoire • orig (texte brut) ; facultatif • expand (texte brut) ; facultatif • subtype (texte brut) ; facultatif • type (valeurs : accord, altvb , inchoative, paradigm) ; obligatoire • xml:lang (code langue de l'IANA) ; facultatif
<code><lang></code>	• du texte brut éventuel	• <code><etym></code>	• norm (valeurs : codes de langue et subtags de variantes de l'IANA) ; obligatoire • orig (texte brut) ; facultatif • expand (texte brut) ; facultatif
<code><langKnowledge></code>	• 0/1/+ <code><lbl></code> • 1/+ <code><langKnown></code>	• <code><person></code>	
<code><langKnown></code>	• du texte brut éventuel	• <code><langKnowledge></code>	• level (valeurs : native...) ; facultatif • tag (code langue de l'IANA) ; obligatoire
<code><location></code>	• 0/1 <code><address></code> • 0/1 <code><country></code> • 0/1 <code><geo></code> • 0/1/+ <code><region></code> • 0/1 <code><settlement></code>	• <code><birth></code> • <code><death></code> • <code><form></code> • <code><residence></code>	
<code><lbl></code>	• du texte brut éventuel	• <code><form></code> • <code><langKnowledge></code> • <code><gramGrp></code> • <code><re></code> • <code><xr></code>	• norm (valeurs : voir en 6.7) ; obligatoire pour le type nature • orig (texte brut) ; facultatif • expand (texte brut) ; facultatif • type (valeurs : catDet, catGen, langKnowledge, nature, reference, subentry) ; obligatoire • xml:lang (code langue de l'IANA) ; obligatoire sous la balise <code><gramGrp></code>
<code><mentioned></code>	• 1 <code><choice></code>	• <code><etym></code>	• xml:id (valeurs : clé unique) ; obligatoire
<code><mood></code>	• rien	• <code><gramGrp></code>	• norm (valeurs : voir en 6.9) ; obligatoire
<code><nameLink></code>	• du texte brut	• <code><nameLink></code>	• sort (valeurs : nombre entier) ; obligatoire
<code><nationality></code>	• du texte brut éventuel	• <code><person></code>	• key (valeurs : codes ISO de pays) ; obligatoire

Balise	Contient	Est contenu dans	Attributs
<code><note></code>	<ul style="list-style-type: none"> du texte brut très rarement, du texte brut dont des passages sont encadrés par <code><hi></code> 	<ul style="list-style-type: none"> <code><bibl></code> <code><cit></code> <code><entry></code> <code><etym></code> <code><form></code> <code><gramGrp></code> <code><persName></code> <code><person></code> <code><re></code> <code><sense></code> <code><xr></code> 	<ul style="list-style-type: none"> place (valeurs : <code>above</code>, <code>below</code>) ; obligatoire rend (valeurs : <code>visible</code>, <code>invisible</code>) ; obligatoire
<code><number></code>	<ul style="list-style-type: none"> rien 	<ul style="list-style-type: none"> <code><gramGrp></code> 	<ul style="list-style-type: none"> norm (valeurs : <code>plural</code>, <code>singular</code>) ; obligatoire
<code><occupation></code>	<ul style="list-style-type: none"> 0/1/+ <code><orgName></code> 0/1/+ <code><roleName></code> 	<ul style="list-style-type: none"> <code><person></code> 	<ul style="list-style-type: none"> from (date) ; facultatif to (date) ; facultatif
<code><orgName></code>	<ul style="list-style-type: none"> du texte brut 	<ul style="list-style-type: none"> <code><occupation></code> 	
<code><orig></code>	<ul style="list-style-type: none"> du texte brut 	<ul style="list-style-type: none"> <code><choice></code> 	
<code><orth></code>	<ul style="list-style-type: none"> du texte brut 	<ul style="list-style-type: none"> <code><form></code> 	<ul style="list-style-type: none"> norm (texte brut) ; obligatoire orig (texte brut) ; facultatif expand (texte brut) ; facultatif value (texte brut) ; obligatoire
<code><per></code>	<ul style="list-style-type: none"> rien 	<ul style="list-style-type: none"> <code><gramGrp></code> 	<ul style="list-style-type: none"> norm (valeurs : 1, 2, 3) ; obligatoire
<code><persName></code>	<ul style="list-style-type: none"> 0/1/+ <code><addName></code> 0/1 <code><choice></code> 0/1/+ <code><forename></code> 0/1/+ <code><genName></code> 0/1/+ <code><namelink></code> 0/1/+ <code><note></code> 0/1/+ <code><roleName></code> 0/1/+ <code><surname></code> 	<ul style="list-style-type: none"> <code><person></code> 	<ul style="list-style-type: none"> xml:lang (code langue de l'IANA) ; facultatif
<code><person></code>	<ul style="list-style-type: none"> 0/1 <code><birth></code> 0/1 <code><death></code> 0/1 <code><gender></code> 0/1/+ <code><idno></code> 0/1 <code><langKnowledge></code> 0/1/+ <code><nationality></code> 0/1/+ <code><occupation></code> 0/1/+ <code><residence></code> 0/1/+ <code><note></code> 1/+ <code><persName></code> 0/1/+ <code><ptr></code> 	<ul style="list-style-type: none"> <code><cit></code> <code><entry></code> <code><re></code> <code><xr></code> 	<ul style="list-style-type: none"> sortkey (valeurs : nom de la personne pour l'ordre alphabétique) xml:id (valeurs : clé unique) ; obligatoire
<code><pos></code>	<ul style="list-style-type: none"> rien 	<ul style="list-style-type: none"> <code><gramGrp></code> 	<ul style="list-style-type: none"> norm (valeurs : voir en 6.12) ; obligatoire
<code><postCode></code>	<ul style="list-style-type: none"> un code postal 	<ul style="list-style-type: none"> <code><address></code> 	

Balise	Contient	Est contenu dans	Attributs
<code><pron></code>	<ul style="list-style-type: none"> du texte brut éventuel 	<ul style="list-style-type: none"> <code><form></code> 	<ul style="list-style-type: none"> extent (valeurs : <code>full</code>, <code>part</code>, <code>poorRhyme</code>, <code>richRhyme</code>, <code>pref</code>, <code>suff</code>) ; obligatoire norm (valeurs diverses) ; obligatoire orig (texte brut) ; facultatif expand (texte brut) ; facultatif notation (valeurs diverses) ; facultatif xml:lang (code langue de l'IANA) ; facultatif
<code><ptr></code>	<ul style="list-style-type: none"> rien 	<ul style="list-style-type: none"> <code><bibl></code> <code><entry></code> <code><form></code> <code><person></code> <code><sense></code> <code><re></code> 	<ul style="list-style-type: none"> subtype (valeurs diverses) ; facultatif target (valeurs : url ou nom de fichier) ; obligatoire targetLang (code langue de l'IANA) ; facultatif type (valeurs : <code>audio</code>, <code>img</code>, <code>uri</code>, <code>url</code>, <code>video</code>, <code>other</code>) ; obligatoire
<code><pubPlace></code>	<ul style="list-style-type: none"> 1 <code><choice></code> 	<ul style="list-style-type: none"> <code><bibl></code> 	<ul style="list-style-type: none"> key (autorité) ; facultatif
<code><publisher></code>	<ul style="list-style-type: none"> 1 <code><choice></code> 	<ul style="list-style-type: none"> <code><bibl></code> 	<ul style="list-style-type: none"> key (autorité) ; facultatif
<code><re></code>	<ul style="list-style-type: none"> 0/1/+ <code><etym></code> 0/1/+ <code><form></code> 0/1/+ <code><note></code> 0/1/+ <code><person></code> 0/1/+ <code><ptr></code> 0/1/+ <code><sense></code> 	<ul style="list-style-type: none"> <code><entry></code> <code><sense></code> 	<ul style="list-style-type: none"> type (valeurs : <code>attestation</code>, <code>citation</code>, <code>collocation</code>, <code>compound</code>, <code>demonym</code>, <code>derived</code>, <code>example</code>, <code>expression</code>, <code>interjection</code>, <code>locution</code>, <code>onomatopeia</code>, <code>proper</code>, <code>pronominal</code>, <code>proverb</code>, <code>unidentified</code>) ; obligatoire xml:id (valeurs : clé unique) ; obligatoire
<code><reg></code>	<ul style="list-style-type: none"> du texte brut 	<ul style="list-style-type: none"> <code><choice></code> 	
<code><region></code>	<ul style="list-style-type: none"> 1 <code><choice></code> 	<ul style="list-style-type: none"> <code><bibl></code> <code><location></code> 	<ul style="list-style-type: none"> key (autorité) ; facultatif type (valeurs : <code>department</code>, <code>region</code>...) ; obligatoire
<code><residence></code>	<ul style="list-style-type: none"> 1 <code><location></code> 	<ul style="list-style-type: none"> <code><person></code> 	<ul style="list-style-type: none"> from (date) ; facultatif to (date) ; facultatif
<code><roleName></code>	<ul style="list-style-type: none"> du texte brut (sous <code><persName></code>) 1 <code><choice></code> (sous <code><occupation></code>) 	<ul style="list-style-type: none"> <code><occupation></code> <code><persName></code> 	<ul style="list-style-type: none"> sort (valeurs : nombre entier) ; obligatoire sous <code><persName></code>
<code><sense></code>	<ul style="list-style-type: none"> 0/1/+ <code><cit></code> 0/1/+ <code><def></code> 0/1/+ <code><note></code> 0/1/+ <code><ptr></code> 0/1/+ <code><re></code> 0/1/+ <code><usg></code> 0/1/+ <code><xr></code> 	<ul style="list-style-type: none"> <code><entry></code> <code><re></code> 	<ul style="list-style-type: none"> level (valeurs : nombre) ; obligatoire n (texte brut) ; facultatif xml:id (valeurs : clé unique) ; obligatoire xml:lang (valeurs : codes de langue et subtags de variantes de l'IANA, voir en 7.1) ; facultatif

Balise	Contient	Est contenu dans	Attributs
<settlement>	• 0/1 <choice>	• <bibl> • <location>	• key (un code commune) ; facultatif • type (city) ; obligatoire
<subc>	• rien	• <gramGrp>	• norm (valeurs : voir en 6.13) ; obligatoire
<surname>	• du texte brut	• <persName>	• type (valeurs : birth, married, paternal, maternal) ; facultatif • sort (valeurs : nombre entier) ; obligatoire
<syll>	• du texte brut éventuel	• <form>	• extent (valeurs : full , part , pref , suff) ; obligatoire • norm ; obligatoire • orig (texte brut) ; facultatif • expand (texte brut) ; facultatif
<title>	• 1 <choice>	• <bibl>	• level (valeurs : a , m , j , s , u) ; obligatoire • key (autorité) ; facultatif • type (valeurs : alt , desc , main , short , sub) ; obligatoire
<tns>	• rien	• <gramGrp>	• norm (valeurs : voir en 6.14) ; obligatoire
<usg>	• du texte brut éventuel	• <cit> • <etym> • <form> • <sense>	• norm (valeurs : voir en 6.15, 6.16, 6.17, 6.18, 6.19 pour les types à contenu normalisé) ; obligatoire • orig (texte brut) ; facultatif • expand (texte brut) ; facultatif • type (ant, binName, colloc, dom, geo, gram, hint, hyper, lang, plev, register, style, synonym, time) ; obligatoire • xml:lang (valeurs : codes de langue et subtags de variantes de l'IANA, voir en 7.1) ; facultatif
<xr>	• 0/1/+ <note> • 0/1/+ <cit> • 1 <lbl> • 0/1/+ <bibl> • 0/1/+ <person>	• <entry> • <sense>	• type (valeurs : abbreviations, antonyms, augmentatives, binName, cf, cognates, contractions, demonyms, deriveds, diminutives, full, hyperonyms, keywords, lemma, sigle, symbol, synonyms, variants) ; obligatoire

1.2 DTD (version technique)

La DTD technique indique un ordre de balises qui n'est pas respecté dans les exemples de ce document et lors de la construction de dictionnaires, afin de faciliter la lisibilité. Un programme remet ensuite toutes les balises dans l'ordre une fois l'ouvrage terminé.

```
<!ELEMENT addName (#PCDATA)>
<!ATTLIST addName sort CDATA #REQUIRED>
<!ATTLIST addName type (epithet | honorific | nick | stagename) #REQUIRED>
```

```

<!ELEMENT address (postCode)>

<!ELEMENT author (choice?)>
<!ATTLIST author key CDATA #IMPLIED>

<!ELEMENT bibl (author*, editor*, date*, title*, edition*, biblScope*, pubPlace*, country*, region*, publisher*, idno*, note*, ptr*)>
<!ATTLIST bibl xml:lang CDATA #IMPLIED>

<!ELEMENT birth (date?, location?)>

<!ELEMENT biblScope (choice)>
<!ATTLIST biblScope unit (chapter | column | issue | line | page | part | verse | volum) #REQUIRED>

<!ELEMENT case (#PCDATA)>
<!ATTLIST case norm (accusative | genitive | nominative | oblique) #REQUIRED>

<!ELEMENT choice (reg, orig?, expan?)>

<!ELEMENT cit (form*, person*, etym*, usg*, note*) >
<!ATTLIST cit type (reference | translation) #REQUIRED>

<!ELEMENT colloc (#PCDATA) >
<!ATTLIST colloc norm CDATA #REQUIRED>
<!ATTLIST colloc orig CDATA #IMPLIED>
<!ATTLIST colloc expand CDATA #IMPLIED>
<!ATTLIST colloc type (pre | post) #REQUIRED>
<!ATTLIST colloc subtype CDATA #IMPLIED>
<!ATTLIST colloc value CDATA #REQUIRED>

<!ELEMENT country (choice?) >
<!ATTLIST country key CDATA #REQUIRED>

<!ELEMENT date (choice?) >
<!ATTLIST date when CDATA #REQUIRED>

<!ELEMENT death (date?, location?)>

<!ELEMENT def (#PCDATA) >
<!ATTLIST def norm CDATA #REQUIRED>
<!ATTLIST def orig CDATA #IMPLIED>
<!ATTLIST def expand CDATA #IMPLIED>
<!ATTLIST def xml:id ID #REQUIRED>
<!ATTLIST def xml:lang CDATA #REQUIRED>

<!ELEMENT edition (choice) >

<!ELEMENT editor (choice?) >
<!ATTLIST editor role CDATA #REQUIRED>
<!ATTLIST editor key CDATA #IMPLIED>

<!ELEMENT entry (ptr*, form*, person*, etym*, sense*, re*, xr*, note*) >
<!ATTLIST entry xml:id ID #REQUIRED>
<!ATTLIST entry cert (confirmed | current | deleted | not_ok | ok) #REQUIRED>

<!ELEMENT etym (lang*, mentioned*, usg*, gloss*, bibl*, note*) >
<!ATTLIST etym corresp CDATA #REQUIRED>

<!ELEMENT expan (#PCDATA) >

```

```

<!ELEMENT forename (#PCDATA)>
<!ATTLIST forename sort CDATA #REQUIRED>

<!ELEMENT form (lbl*, lang*, colloc*, orth, syll*, pron*, form*, colloc*, gramGrp*, bibl*, usg*, ptr*, location*, note*)>
<!ATTLIST form cert (current | not_ok | not_ok_var | ok) #REQUIRED>
<!ATTLIST form n CDATA #IMPLIED>
<!ATTLIST form subtype (morpheme | phrase | symbol | word) #REQUIRED>
<!ATTLIST form type (elided | inflected | inflectelided | lemma) #REQUIRED>
<!ATTLIST form value CDATA #REQUIRED>
<!ATTLIST form xml:id ID #REQUIRED>
<!ATTLIST form xml:lang CDATA #REQUIRED>

<!ELEMENT gen (#PCDATA) >
<!ATTLIST gen norm (feminine | masculine | neuter) #REQUIRED>

<!ELEMENT gender (choice?)>
<!ATTLIST gender value (feminine | masculine | other) #REQUIRED>

<!ELEMENT genName (#PCDATA)>
<!ATTLIST genName sort CDATA #REQUIRED>

<!ELEMENT geo (#PCDATA) >

<!ELEMENT gloss (choice) >
<!ATTLIST gloss corresp CDATA #IMPLIED>
<!ATTLIST gloss type (definition | transcription | translation | type | divers) #REQUIRED>

<!ELEMENT gram (#PCDATA) >
<!ATTLIST gram norm CDATA #REQUIRED>
<!ATTLIST gram type (catDet | catFam | catGen | catEagles | catGrace | polarity) #REQUIRED>

<!ELEMENT gramGrp (pos?, subc?, per?, gen?, number?, tns?, mood?, case?, gram*, iType*, lbl*, note*) >

<!ELEMENT hi (#PCDATA) >
<!ATTLIST hi style CDATA #REQUIRED>

<!ELEMENT idno (#PCDATA) >
<!ATTLIST idno type CDATA #REQUIRED>

<!ELEMENT iType (#PCDATA) >
<!ATTLIST iType norm CDATA #REQUIRED>
<!ATTLIST iType orig CDATA #IMPLIED>
<!ATTLIST iType expand CDATA #IMPLIED>
<!ATTLIST iType subtype CDATA #IMPLIED>
<!ATTLIST iType type (accord | altvb | inchoative | paradigm) #REQUIRED>
<!ATTLIST iType xml:lang CDATA #IMPLIED>

<!ELEMENT lang (#PCDATA) >
<!ATTLIST lang norm CDATA #REQUIRED>
<!ATTLIST lang orig CDATA #IMPLIED>
<!ATTLIST lang expand CDATA #IMPLIED>

<!ELEMENT langKnowledge (lbl?, langKnown+) >

<!ELEMENT langKnown (choice?)>
<!ATTLIST langKnown level CDATA #IMPLIED>
<!ATTLIST langKnown tag CDATA #REQUIRED>

```

```

<!ELEMENT lbl (#PCDATA) >
<!ATTLIST lbl norm CDATA #IMPLIED>
<!ATTLIST lbl orig CDATA #IMPLIED>
<!ATTLIST lbl expand CDATA #IMPLIED>
<!ATTLIST lbl type (catDet | catGen | langKnowledge | nature | reference | subentry) #REQUIRED>
<!ATTLIST lbl xml:lang CDATA #IMPLIED>

<!ELEMENT location (address?, settlement?, region*, country?, geo?) >

<!ELEMENT mentioned (choice) >
<!ATTLIST mentioned xml:id ID #REQUIRED>

<!ELEMENT mood (#PCDATA) >
<!ATTLIST mood norm (conditional | gerundive | imperative | indicative | infinitive | participle | subjunctive ) #REQUIRED>

<!ELEMENT nameLink (#PCDATA)>
<!ATTLIST nameLink sort CDATA #REQUIRED>

<!ELEMENT nationality (choice?) >
<!ATTLIST nationality key CDATA #REQUIRED>

<!ELEMENT note (#PCDATA | hi)* >
<!ATTLIST note place (above | below) #REQUIRED>
<!ATTLIST note rend (invisible | visible) #REQUIRED>

<!ELEMENT number (#PCDATA) >
<!ATTLIST number norm (plural | singular) #REQUIRED>

<!ELEMENT occupation (roleName?, orgName?) >
<!ATTLIST occupation from CDATA #IMPLIED>
<!ATTLIST occupation to CDATA #IMPLIED>

<!ELEMENT orgName (choice) >

<!ELEMENT orig (#PCDATA) >

<!ELEMENT orth (#PCDATA)* >
<!ATTLIST orth norm CDATA #REQUIRED>
<!ATTLIST orth orig CDATA #IMPLIED>
<!ATTLIST orth expand CDATA #IMPLIED>
<!ATTLIST orth value CDATA #REQUIRED>

<!ELEMENT per (#PCDATA) >
<!ATTLIST per norm (1 | 2 | 3) #REQUIRED>

<!ELEMENT persName (addName*, roleName*, forename*, namelink*, genName*, addName*, choice*, note*) >
<!ATTLIST persName xml:lang CDATA #IMPLIED>

<!ELEMENT person (persName+, birth?, death?, gender?, occupation*, residence*, nationality*, idno*, langKnowledge?, ptr*, note*) >
<!ATTLIST person sortkey CDATA #REQUIRED>
<!ATTLIST person xml:id ID #REQUIRED>

<!ELEMENT pos (#PCDATA) >
<!ATTLIST pos norm CDATA #REQUIRED>

<!ELEMENT postCode (#PCDATA) >

```

```

<!ELEMENT pron (#PCDATA) >
<!ATTLIST pron extent (full | part | poorRhyme | richRhyme | pref | suff) #REQUIRED>
<!ATTLIST pron norm CDATA #REQUIRED>
<!ATTLIST pron orig CDATA #IMPLIED>
<!ATTLIST pron expand CDATA #IMPLIED>
<!ATTLIST pron notation CDATA #IMPLIED>
<!ATTLIST pron xml:lang CDATA #REQUIRED>

<!ELEMENT ptr (#PCDATA) >
<!ATTLIST ptr target CDATA #REQUIRED>
<!ATTLIST ptr subtype CDATA #IMPLIED>
<!ATTLIST ptr type (audio | img | uri | url | video | other) #REQUIRED>
<!ATTLIST ptr targetLang CDATA #IMPLIED>

<!ELEMENT pubPlace (choice?) >
<!ATTLIST pubPlace key CDATA #IMPLIED>

<!ELEMENT publisher (choice?) >
<!ATTLIST publisher key CDATA #IMPLIED>

<!ELEMENT re (lbl?, form*, person*, etym*, sense*, ptr*, note*) >
<!ATTLIST re type (attestation | citation | collocation | compound | demonym | derived | example | expression
| interjection | locution | onomatopeia | pronominal | proper | proverb | unidentified) #REQUIRED>
<!ATTLIST re xml:id ID #REQUIRED>

<!ELEMENT reg (#PCDATA) >

<!ELEMENT region (choice?) >
<!ATTLIST region key CDATA #IMPLIED>
<!ATTLIST region type CDATA #REQUIRED>

<!ELEMENT residence (location) >
<!ATTLIST residence from CDATA #IMPLIED>
<!ATTLIST residence to CDATA #IMPLIED>

<!ELEMENT roleName (#PCDATA | choice)*>
<!ATTLIST roleName sort CDATA #IMPLIED>

<!ELEMENT sense (usg*, lang*, def*, cit*, note*, xr*, re*, ptr*) >
<!ATTLIST sense level CDATA #REQUIRED>
<!ATTLIST sense n CDATA #IMPLIED>
<!ATTLIST sense xml:id ID #REQUIRED>
<!ATTLIST sense xml:lang CDATA #IMPLIED>

<!ELEMENT settlement (choice?) >
<!ATTLIST settlement key CDATA #IMPLIED>
<!ATTLIST settlement type (city) #REQUIRED>

<!ELEMENT surname (#PCDATA)>
<!ATTLIST surname sort CDATA #REQUIRED>
<!ATTLIST surname type (birth | married | maternal | paternal) #REQUIRED>

<!ELEMENT subc (#PCDATA) >
<!ATTLIST subc norm CDATA #REQUIRED>

<!ELEMENT syll (#PCDATA) >
<!ATTLIST syll extent (full | part | pref | suff) #REQUIRED>
<!ATTLIST syll norm CDATA #REQUIRED>
<!ATTLIST syll orig CDATA #IMPLIED>
<!ATTLIST syll expand CDATA #IMPLIED>

```

```
<!ELEMENT title (choice?) >
<!ATTLIST title level (a | m | j | s | u) #REQUIRED>
<!ATTLIST title type (alt | desc | main | short | sub) #REQUIRED>
<!ATTLIST title key CDATA #IMPLIED>

<!ELEMENT tns (#PCDATA) >
<!ATTLIST tns norm CDATA #REQUIRED>

<!ELEMENT usg (#PCDATA) >
<!ATTLIST usg norm CDATA #REQUIRED>
<!ATTLIST usg orig CDATA #IMPLIED>
<!ATTLIST usg expand CDATA #IMPLIED>
<!ATTLIST usg type (ant | binName | colloc | dom | geo | gram | hint | hyper | lang | lemma | plev | register |
style | synonym | time) #REQUIRED>
<!ATTLIST usg xml:lang CDATA #IMPLIED>

<!ELEMENT xr (lbl, person*, bibl*, cit*, note*) >
<!ATTLIST xr type (abbreviations | antonyms | augmentatives | binName | cf | cognates | contractions |
demonyms | deriveds | diminutives | full | hyperonyms | keywords | lemma | sigle | symbol | synonyms |
variants) #REQUIRED>
```

2 Structure générale des entrées

Chaque entrée **<entry>** a une clé d'identification (« **xml:id** ») de type "entry-[nombre]" qui lui est propre. L'ordre de ses éléments se présente généralement comme suit, mais il n'est pas immuable :

On trouve en premier dans une entrée la forme du mot-vedette dans la langue du dictionnaire. Il peut, en cas de variantes ou de synonymes, y avoir plusieurs formes qui se suivent. Chaque forme contient l'orthographe, la ou les prononciations, un éventuel découpage en syllabe, les informations grammaticales, les éventuelles collocations, les éventuelles formes fléchies (féminin, pluriel, forme conjuguée) et élidées, certaines informations d'usage (registre notamment), de nature (abréviation, contraction...), de localisation (pays, région, coordonnées pour les formes qui sont des lieux), d'éventuels fichiers liés (illustration, enregistrement...), des références bibliographiques et notes éventuelles.

Si l'entrée est une personne et pas un mot, à la place du mot-vedette, on a une « personne-vedette ».

On peut trouver ensuite les informations d'étymologie avec la langue d'origine, le terme d'origine, éventuellement des explications sur ce terme, des informations d'usage, de date, de bibliographie.

On trouve enfin les différentes acceptances du mot, numérotées.

Chaque acceptance contient en général une ou des définitions, ou bien une ou des traductions contenant elles-mêmes une ou des formes (termes). La forme peut être complétée par des informations étymologiques, des informations d'usage, grammaticales, de nature ou des notes.

L'acceptation contient également des sous-entrées (expressions, exemples, citations, mots composés, dérivés...). Elles se présentent comme des entrées : forme, informations étymologiques, acceptances.

L'acceptation peut enfin contenir des informations d'usage, des notes ou des références (vers d'autres mots du dictionnaire, vers des mots ne se trouvant pas dans le dictionnaire, vers un ouvrage : renvois, synonymes, variantes, mots-clés...).

3 Structure détaillée des entrées

3.1 L'entrée

La balise **<entry>** est complétée par l'attribut « **xml:id** ». Il contient la clé d'identification de l'entrée (**entry-[nombre]**). Elle permet d'ordonner les entrées de manière pertinente.

Elle peut également avoir un attribut « **cert** », notamment dans les documents en cours de construction. Il sert à indiquer l'état de l'entrée et prend les valeurs suivantes :

- **current** : entrée en cours de modification
- **deleted** : entrée à supprimer
- **not_ok** : entrée pas encore validée
- **ok** : entrée validée
- **confirmed** : entrée pré-validée par ses rédacteurs, mais pas par l'équipe de relecture

Exemple :

```
balise <entry>
<entry xml:id="entry-1805" cert="ok">
...
</entry>
```

3.2 Le mot vedette

Le mot vedette se trouve entre les balises **<form>** et **</form>**. Il peut y avoir plusieurs mots vedettes côte à côte ou imbriqués.

La balise **<form>** comporte un attribut « **xml:lang** » indiquant sa langue et sa variété grâce aux codes langue de l'IANA. (voir 7.1).

Exemple :

```
« maison »
<form subtype="word" type="lemma" xml:lang="fr" xml:id="form-11" cert="ok" n="1" value="maison">
...
</form>
```

Lorsqu'il s'agit d'une forme multilingue ou multivariété, on les indique toutes dans l'attribut « **xml:lang** » en les séparant par « / ».

Exemple :

```
« ostau »
<form subtype="word" type="lemma" xml:lang="oc-auvern-grclass/oc-gascon-grclass/oc-provenc-
grclass/oc-vivaraup-grclass" xml:id="form-12" cert="ok" n="1" value="ostau">
...
</form>
```

La forme a également un attribut « **xml:id** » qui est un identifiant unique pour les formes d'un document. Il prend la forme **form-[nombre]**.

L'attribut « **type** » permet d'indiquer le niveau de la forme. Il prend les valeurs suivantes :

- **elided** : forme élidée (de la forme parente ou de la forme enfant)
- **inflected** : forme fléchie (de la forme parente ou de la forme enfant)
- **inflectelidied** : forme fléchie élidée (de la forme parente ou de la forme enfant)
- **lemma** : lemme (de la forme parente ou de la forme enfant)

Dans la plupart des lexiques, on a comme forme de premier niveau le lemme et comme forme enfant les éventuelles formes fléchies. Mais dans certains lexiques, les formes fléchies sont au premier niveau et on a besoin de connaître leur lemme pour les différencier les unes des autres (ex. : « *vòli* » forme fléchie de « *volar* » et « *vòli* » forme fléchie de « *voler* »). Dans ces cas-là, on peut avoir un **<form type="inflected">** en premier niveau avec un **<form type="lemma">** comme enfant.

La balise **<form>** est complétée par un attribut « **subtype** » qui prend les valeurs suivantes :

- **morpheme** : si la forme est un suffixe, un préfixe, une lettre...
- **phrase** : item lexical qui n'est pas une unité grammaticale ni une partie de mot, mais plusieurs formes rassemblées. C'est une phrase ou un bout de phrase (phrases d'exemples, certaines expressions imagées, etc.)
- **symbol** : si la forme est un symbole, une unité de mesure, etc.
- **word** : unité grammaticale.

Pour bien comprendre la distinction entre unité grammaticale et bout de phrase, voir en 7.2.

Exemples :

« *opération* »

```
<form subtype="word" type="lemma" xml:lang="fr" xml:id="form-124" cert="ok" n="1" value="opération">
  <orth norm="opération" value="opération">opération</orth>
  ...
</form>
```

« *canalha, canalhas* »

```
<form subtype="word" type="lemma" xml:lang="oc-lengadoc-grclass" xml:id="form-14" n="1"
  value="canalha">
  <orth norm="canalha" value="canalha">canalha</orth>
  <form subtype="word" type="inflected" xml:lang="oc-lengadoc-grclass" xml:id="form-15" n="1"
  value="canalhas">
    <orth norm="canalhas" value="canalhas">canalhas</orth>
    ...
  </form>
  ...
</form>
```

La balise **<form>** a un attribut « **value** » qui indique sa transcription incluant les collocations et l'orthographe dans l'ordre, traitée pour la recherche (remplacement des æ et des œ par des ae et des oe, remplacement des - par des espaces, les ' par des ', des ponctuations par des espaces, suppression des espaces en début et fin de chaîne, suppression des doubles espaces).

Exemple :

« *se parler de* »

```

<form subtype="word" type="lemma" xml:lang="fr" xml:id="form-435" cert="ok" value="se parler de">
  <colloc type="pre" norm="se " orig="se " value="se">se </colloc>
  <orth norm="parler" value="parler" orig="parler">parler</orth>
  <colloc type="pre" norm=" de" orig=" de" value="de"> de</colloc>
  ...
</form>

```

La balise **<form>** a également un attribut « **cert** » qui permet d'indiquer si la forme est préconisée ou contestée par le *Congrès*. Il prend les valeurs suivantes :

- **ok** : forme validée par le *Congrès*
- **not_ok** : forme contestée par le *Congrès*
- **not_ok_var** : forme contestée par le *Congrès* dans certaines variétés
- **current** : forme ni validée, ni contestée (= encore en cours d'examen)

Exemple :

<i>Original</i> : « <i>normalizar</i> ; * <i>normalisar</i> »
<i>Affiché</i> (éd. papier et numérique) : « <i>normalizar</i> ; * <i>normalisar</i> »
<form subtype="word" type="lemma" xml:lang="oc-lengadoc-grclass" xml:id="form-756" cert="ok" value="normalizar"> <orth norm="normalizar" value="normalizar" orig="normalizar">normalizar</orth> ... </form> <form subtype="word" type="lemma" xml:lang="oc-lengadoc-grclass" xml:id="form-757" cert="not_ok" value="normalisar"> <orth norm="normalisar" value="normalisar" orig="normalisar">normalisar</orth> ... </form>

La balise **<form>** est en général complétée par un attribut « **n** » qui est son identifiant de référence dans le cas d'homographes. Il contient un numéro qui servira à différencier la forme de chaque homographe.

Il sert notamment pour les formes fléchies de deux lemmes différents qui ont la même orthographe et la même catégorie grammaticale. Par exemple, « *vòli* » peut être la première personne du singulier au présent de l'indicatif de « *voler* » (*vouloir*) ou de « *volar* » (*voler*).

Il sert aussi pour les lemmes qui se fléchissent différemment mais qui ont la même orthographe et la même catégorie grammaticale. Par exemple, les verbes « *estar* » en gascon dont l'un se conjugue « *soi, es, esto...* » et l'autre « *estau, està, estei* ».

L'identifiant homographe des mots qui n'ont pas d'homographe est « *1* ».

Les **<form subtype="phrase">** qui n'ont pas de catégorie détaillée n'ont pas d'identifiant homographe.

Exemples :

« <i>desparéisher, despareish ... desparir, despareish</i> »
<entry xml:id="entry-12"> <form subtype="word" type="lemma" xml:lang="oc-gascon-grclass" n="1" xml:id="form-28" cert="ok" value="desparéisher"> <orth norm="desparéisher" value="desparéisher">desparéisher</orth> ... </form>

```

<form subtype="word" type="inflected" xml:lang="oc-gascon-grclass" xml:id="form-29" n="1" cert="ok"
value="despareish">
<orth norm="despareish" value="despareish">despareish</orth>
<gramGrp>
  <gram type="catDet" norm="VerbeIndPres1s"></gram>
  <gram type="catGen" norm="Verbe"></gram>
</gramGrp>
...
</form>
<gramGrp>
  <gram type="catDet" norm="VerbeInf"></gram>
  <gram type="catGen" norm="Verbe"></gram>
</gramGrp>
...
</form>
</entry>

<entry xml:id="entry-20">
<form subtype="word" type="lemma" xml:lang="oc-gascon-grclass" n="1" xml:id="form-51" cert="ok"
value="desparir">
<orth norm="desparir" value="desparir">desparir</orth>
<form subtype="word" type="inflected" xml:lang="oc-gascon-grclass" xml:id="forme-52" n="2" cert="ok"
value="despareish">
  <orth norm="despareish" value="despareish">despareish</orth>
  <gramGrp>
    <gram type="catDet" norm="VerbeIndPres1s"></gram>
    <gram type="catGen" norm="Verbe"></gram>
  </gramGrp>
  ...
</form>
<gramGrp>
  <gram type="catDet" norm="VerbeInf"></gram>
  <gram type="catGen" norm="Verbe"></gram>
</gramGrp>
...
</form>
</entry>

```

« abocar (o) / abocar (ò) »

```

<entry xml:id="entry-50">
<form subtype="word" type="lemma" xml:lang="oc-lengadoc-grclass" n="1" xml:id="form-25" cert="ok"
value="abocar">
  <orth norm="abocar" value="abocar">abocar</orth>
  <gramGrp>
    <gram type="catDet" norm="VerbeInf"></gram>
    <gram type="catGen" norm="Verbe"></gram>
  </gramGrp>
  <gramGrp>
    <iType type="altvb" norm="o/o" orig="o">o</iType>
  </gramGrp>
  ...
</form>
</entry>

<entry xml:id="51">
<form subtype="word" type="lemma" xml:lang="oc-lengadoc-grclass" n="2" xml:id="form-26" cert="ok"
value="abocar">
  <orth norm="abocar" value="abocar">abocar</orth>

```

```

<gramGrp>
  <gram type="catDet" norm="VerbeInf"></gram>
  <gram type="catGen" norm="Verbe"></gram>
</gramGrp>
<gramGrp>
  <iType type="altvb" norm="ò/o" orig="ò">ò</iType>
</gramGrp>
...
</form>
</entry>

```

3.2.1 La nature

La balise **<lbl type="nature">** est utilisée pour donner une information sur la nature de la forme (abréviation, contraction... cf 6.7) et/ou sur ses relations avec une autre forme (synonyme, variante).

Elle porte les attributs « **norm** », « **orig** » et « **expand** ». (cf. 4).

Exemples :

<i>Original : « BD (abréviation de Bande dessinée) »</i>
<i>Affiché (éd. papier) : « BD (abréviation) / bande dessinée »</i>
<i>Affiché (éd. numérique) : « BD (abréviation) / bande dessinée »</i>
<form subtype="word" type="lemma" xml:lang="fr" xml:id="form-5" n="1" cert="ok" value="BD"> <orth norm="BD" value="BD" orig="BD">BD</orth> ... <lbl type="nature" norm="abbreviation" orig="abréviation de Bande dessinée">abréviation</lbl> </form> <form subtype="phrase" type="lemma" xml:lang="fr" xml:id="form-6" n="1" cert="ok" value="bande dessinée"> <orth norm="bande dessinée" value="bande dessinée">bande dessinée</orth> ... </form>

<i>Original : « ampère-òra / symbole : Ah »</i>
<i>Affiché (éd. papier et numérique) : « ampère-òra / Ah (symbole) »</i>
<form subtype="word" type="lemma" xml:lang="oc-gascon-grclass" xml:id="form-20" n="1" value="ampère òra"> <orth norm="ampère-òra" value="ampère-òra" orig="ampère òra">ampère-òra</orth> ... </form> <form subtype="word" type="lemma" xml:lang="oc-gascon-grclass" xml:id="form-21" n="1" value="Ah"> <lbl type="nature" norm="symbol" orig="symbole">symbole</lbl> <orth norm="Ah" value="Ah" orig="Ah">Ah</orth> ... </form>

<i>Original : « bipeur / biper »</i>
<i>Affiché (éd. papier et numérique) : « bipeur / biper »</i>

```

<form subtype="word" type="lemma" xml:lang="fr" xml:id="form-3" n="1" value="bipeur">
  <orth norm="bipeur" value="bipeur" orig="bipeur">bipeur</orth>
  ...
</form>
<form subtype="word" type="lemma" xml:lang="fr" xml:id="form-4" n="1" value="biper">
  <orth norm="biper" value="biper" orig="biper">biper</orth>
  <lbl type="nature" norm="variant" orig="var."></lbl>
  ...
</form>

```

auto / automobile

```

<form subtype="word" type="lemma" xml:lang="fr" xml:id="form-203" n="1" cert="current" value="auto">
  <lbl type="nature" norm="abbreviation"></lbl>
  <orth norm="auto" value="auto" orig="auto">auto</orth>
  ...
</form>
<form subtype="word" type="lemma" xml:lang="fr" xml:id="form-204" n="1" cert="current"
value="automobile">
  <orth norm="automobile" value="automobile" orig="automobile">automobile</orth>
  ...
</form>

```

3.2.2 L'orthographe

Elle est placée entre les balises **<orth>** et **</orth>**. Elle est obligatoire, mais peut être invisible si sa valeur est contenue dans l'attribut « **norm** » et que le contenu entre balises est laissé vide.

La balise **<orth>** porte les attributs « **norm** », « **orig** » et « **expand** » (cf. 4). Elle comporte aussi, dans la plupart des dictionnaires, un attribut « **value** » dont la valeur est une orthographe traitée pour la recherche (remplacement des æ et des œ par des ae et des oe, remplacement des - par des espaces, les ' par des ', des ponctuations par des espaces, suppression des doubles espaces et des espaces en début et fin de chaîne).

Exemples :

Original : « manger »

Affiché (éd. papier) : « manger »

Affiché (éd. numérique) : « manger »

```

<form subtype="word" type="lemma" xml:lang="fr" xml:id="form-312" n="1" cert="ok" value="manger">
  <orth norm="manger" value="manger" orig="manger">manger</orth>
  ...
</form>

```

Original : « oeil-de-boeuf »

Affiché (éd. papier) : « œil-de-bœuf »

Affiché (éd. numérique) : « œil-de-bœuf »

```

<form subtype="word" type="lemma" xml:lang="fr" xml:id="form-3" n="1" cert="ok" value="oeil de boeuf">
  <orth norm="œil-de-bœuf" value="oeil de boeuf" orig="œil-de-boeuf">œil-de-bœuf</orth>

```

```
...  
</form>
```

Affiché (éd. papier et numérique) : « remplacer par le féminin »

```
<note rend="visible" place="above">remplacer par le féminin</note>  
<form subtype="word" type="lemma" xml:lang="fr" xml:id="form-200" n="1" cert="ok" value="grande">  
  <orth norm="grande" value="grande"></orth>  
  ...  
</form>
```

3.2.3 Les informations bibliographiques

Si une forme particulière est préconisée par un auteur, dans un ouvrage..., on insère les balises **<bibl>** et **</bibl>** dans la forme pour indiquer les références.

Si les informations bibliographiques ne concernent que certaines variétés d'une forme qui existe dans plusieurs variétés, on peut l'indiquer grâce à l'attribut « **xml:lang** ». Il prend en valeur un ou plusieurs codes langues de l'IANA (voir en 7.1) séparés par « / ».

Affiché (éd. papier et numérique) : « TDF (*Frederic Mistral, 1878*) »

```
<bibl xml:lang="oc-lengadoc-grclass">  
  <title type="short" level="m">  
    <choice>  
      <reg>Dic. Laus</reg>  
    </choice>  
  </title>  
</bibl>  
<bibl xml:lang="oc-gascon-grclass">  
  <title type="short" level="m">  
    <choice>  
      <reg>Dic. Per Noste</reg>  
    </choice>  
  </title>  
</bibl>  
<bibl xml:lang="oc-lengadoc-grclass/oc-gascon-grclass">  
  <title type="short" level="m">  
    <choice>  
      <reg>Basic</reg>  
    </choice>  
  </title>  
</bibl>
```

La balise **<bibl>** peut contenir les balises suivantes :

- **<author>** : auteur
- **<biblScope>** : l'extension de l'œuvre, par exemple les numéros de page
- **<country>** : le pays de publication
- **<date>** : date de publication
- **<edition>** : le numéro d'édition de l'œuvre (première édition, deuxième édition...)
- **<editor>** : personne ayant une responsabilité secondaire (éditeur scientifique, compilateur, traducteur...).
- **<idno>** : l'identifiant de référence du livre (ISBN, ISSN, DOI...)
- **<note>** : informations diverses

- **<publisher>** : éditeur
- **<pubPlace>** : le lieu de publication
- **<ptr>** : les fichiers liés
- **<region>** : la région, le département de publication
- **<title>** : titre de l'œuvre

Pour indiquer les valeurs de chacun de ces éléments (sauf **<note>** et **<idno>**), on utilise les balises **<reg>**, **<orig>** et **<expan>** (voir 4).

La balise **<author>** peut être complétée par un attribut « **key** » renvoyant à l'autorité de la personne à laquelle elle fait référence (son prénom et son nom complet, tels qu'enregistrés dans l'Etat civil), notamment si le nom entre les balises n'est pas le nom complet.

La balise **<biblScope>** est complétée par un attribut « **unit** » pouvant prendre les valeurs suivantes :

- **chapter** : numéro et/ou titre de chapitre
- **column** : colonne
- **issue** : numéro de livraison
- **line** : numéro de ligne
- **page** : numéro ou étendue de pages
- **part** : autres
- **verse** : numéro du vers
- **volum** : numéro de volume

La balise **<country>** est complétée par un attribut « **key** » qui contient un code ISO 3166 de pays.

La balise **<date>** est complétée par un attribut « **when** » qui contient une date à un format normalisé (de type AAAA ou AAAA-MM ou AAAA-MM-JJ). Quand il s'agit d'un siècle, on met la première année du siècle.

La balise **<editor>** est complétée par un attribut « **role** » pouvant prendre les valeurs suivantes :

- **compiler** : compilateur
- **illustrator** : illustrateur
- **translator** : traducteur

Elle peut également être complétée par un attribut « **key** » renvoyant à l'autorité de la personne à laquelle elle fait référence (son prénom et son nom complet, tels qu'enregistrés dans l'Etat civil), notamment si le nom entre les balises n'est pas le nom complet.

La balise **<idno>** est complétée par un attribut « **type** » qui permet d'indiquer la nature de l'identifiant : **ISBN**, **ISSN**, **DOI**, **QID**, **URI**, **VIAF**, **ESTC**, **OCLC**.

La balise **<ptr>** sert à indiquer un fichier lié (image de la couverture, URL pour acheter l'ouvrage...). Voir en 3.6.

Les balises **<publisher>** et **<pubPlace>** peuvent être complétées par un attribut « **key** » renvoyant à l'autorité respectivement de la structure et du lieu auxquelles elles font référence, notamment si le nom entre les balises n'est pas le nom complet.

La balise **<region>** est complétée par un attribut « **type** » qui indique le type de subdivision géographique (**department**, **region**...), et, éventuellement, par un attribut « **key** » renvoyant à l'autorité de la zone concernée (son nom officiel).

La balise **<title>** est complétée par un attribut « **type** » pouvant prendre les valeurs suivantes :

- **alt** : titre alternatif (par exemple dans une autre langue) par lequel l'ouvrage est connu
- **desc** : paraphrase descriptive de l'œuvre ayant les fonctions d'un titre
- **main** : titre principal
- **short** : forme abrégée d'un titre
- **sub** : sous-titre, titre de partie

Elle est également complétée par un attribut « **level** » qui permet de citer le niveau et la nature du texte qu'il intitule. Il prend les valeurs suivantes :

- **a** : titre analytique, pour des textes publiés comme parties d'un ensemble plus grand (article, poème...)
- **j** : titre de journal, de revue
- **m** : titre de monographie, pour des textes publiés comme un document distinct (livre, ensemble, volume d'un ouvrage...)
- **s** : titre de publication en série (bulletins...)
- **u** : titre de matière non publié (thèses, dissertations non éditées commercialement...)

Si le titre n'est pas écrit sous sa forme complète, on peut utiliser l'attribut « **key** » pour indiquer celle-ci.

Exemples :

<i>Original : « préconisé dans le Tresór de Mistral »</i>
<i>Affiché (éd. papier) : « TDF (Frederic Mistral, 1878) »</i>
<i>Affiché (éd. numérique) : « Tres. du Fel. (Frederic Mistral, 1878) »</i>
<pre><bibl> <note rend="invisible" place="above">préconisé dans le</note> <title type="short" level="m" key="Tresór dou Felibrige"> <choice> <orig>Tresór</orig> <reg>TDF</reg> <expan>Tres. du Fel.</expan> </choice> </title> <author key="Frédéric Mistral"> <choice> <orig>Mistral</orig> <reg>Frederic Mistral</reg> </choice> </author> <date when="1878"> <choice> <reg>1878</reg> </choice> </date> </bibl></pre>

Affiché (éd. papier et numérique) : « Dictionnaire français-occitan, Languedocien central ; *Laus* (97) ; Section du tarn de l'IEO : Albi ; p.141 »

```
<bibl>
<title type="main" level="m">
<choice>
<reg>Dictionnaire français-occitan</reg>
</choice>
</title>
<title type="sub" level="m">
<choice>
<reg>Languedocien central</reg>
</choice>
</title>
<author key="Christian Laux">
<choice>
<reg>Laus</reg>
</choice>
</author>
<date when="1997">
<choice>
<reg>97</reg>
</choice>
</date>
<publisher key="IEO Tarn">
<choice>
<reg>Section du Tarn de l'IEO</reg>
</choice>
</publisher>
<pubPlace>
<choice>
<reg>Albi</reg>
</choice>
</pubPlace>
<biblScope type="page">
<choice>
<reg>p. 141</reg>
</choice>
</biblScope>
</bibl>
```

Affiché (éd. papier et numérique) : (*Camps Christian, 2008. Expressions et dictons occitans, in Lengas 63*) »

```
<bibl>
<author key="Christian Camps">
<choice>
<reg>Camps Christian</reg>
</choice>
</author>
<date when="2008">
<choice>
<reg>2008</reg>
</choice>
</date>
<title type="main" level="a">
<choice>
<reg>Expressions et dictons occitans</reg>
```

```

</choice>
</title>
<title type="main" level="j">
  <choice>
    <reg>Lengas</reg>
  </choice>
</title>
<biblScope unit="volum">
  <choice>
    <reg>63</reg>
  </choice>
</title>
</bibl>

```

Exemple d'utilisation des balises `<bibl>` pour indiquer une préconisation :

Original : « *grama* (unité) l (*Alibèrt*) *gram* »

Affiché (éd. papier et numérique) : « (*unité*) *grama* l (*Alibèrt*) *gram* »

```

<cit type="translation">
  <form subtype="word" type="lemma" xml:lang="oc-lengadoc-grclass" xml:id="form-4" n="1" cert="current"
value="grama">
    <orth norm="grama" value="grama" orig="grama">grama</orth>
    ...
  </form>
  <usg type="hyper" norm="unité" orig="unité">unité</usg>
  <form subtype="word" type="lemma" xml:lang="oc-lengadoc-grclass" xml:id="form-5" n="1" cert="current"
value="gram">
    <bibl>
      <author key="Louis Alibert">
        <orig>Alibèrt</orig>
        <reg>Alibèrt</reg>
      </author>
    </bibl>
    <orth norm="gram" value="gram" orig="gram">gram</orth>
    ...
  </form>
</cit>

```

3.2.4 Les collocations

Elles sont placées entre les balises `<colloc>` et `</colloc>`, complétées par l'attribut « `type` » qui permet d'indiquer la position de la collocation. Il prend les valeurs suivantes :

- **pre** : avant le mot-vedette
- **post** : après le mot-vedette

L'attribut « `subtype` » peut éventuellement être utilisé pour préciser la catégorie grammaticale de la collocation en utilisant un code de catégorie détaillée (5.1).

La balise `<colloc>` porte les attributs « `norm` », « `orig` » et « `expand` » (cf. 4).

Le contenu entre les balises et celui des attributs « `norm` », « `orig` » et « `expand` » contient l'espace qui sépare la collocation du mot lorsqu'il y en a un. La collocation sans espace apparaît dans l'attribut « `value` », construit comme celui de l'orthographe (voir en 3.2.2).

Exemples :

Affiché (éd. papier et numérique) : « *se targuer (de)* »

```
<form subtype="word" type="lemma" xml:lang="fr" xml:id="form-12" n="1" cert="ok" value="se targuer de">
<colloc type="pre" norm="se " value="se">se </colloc>
<orth norm="targuer" value="targuer">targuer</orth>
<colloc type="post" norm=" de" value="de"> de</colloc>
...
</form>
```

Affiché (éd. papier et numérique) : « *(se) dire* »

```
<form subtype="word" type="lemma" xml:lang="fr" xml:id="form-59" n="1" cert="ok" value="se dire">
<orth norm="dire" value="dire">dire</orth>
<colloc type="pre" subtype="Prep" norm="se " value="se">se </colloc>
</form>
```

Les collocations permettent de retrouver une forme, même si elle est utilisée avec un autre mot qui n'est pas important. Par exemple, elle permet de retrouver « parler de » même si l'utilisateur recherche uniquement le mot « parler ». Par contre, on ne considère pas comme collocation un mot faisant pleinement partie d'une locution. Par exemple, on ne retrouvera pas « par rapport à » en recherchant « rapport ».

Exemples :

Original : « *les bovins* »

Affiché (éd. papier et numérique) : « *(les) bovins* »

```
<form subtype="word" type="lemma" xml:lang="fr" xml:id="form-10" n="1" cert="current" value="les bovins">
<colloc type="pre" norm="les " orig="les " value="les">les </colloc>
<orth norm="bovins" value="bovins" orig="bovins">bovins</orth>
...
</form>
```

Original : « *les doigts de pied en éventail* »

Affiché (éd. papier et numérique) : « *les doigts de pied en éventail* »

```
<form subtype="phrase" type="lemma" xml:lang="fr" xml:id="form-4510" n="1" cert="current" value="les doigts de pied en éventail">
<orth norm="les doigts de pied en éventail" value="les doigts de pied en éventail" orig="les doigts de pied en éventail">les doigts de pied en éventail</orth>
...
</form>
```

Si une forme multivariété a des collocations qui n'existent que dans certaines variétés, on ne peut pas mettre ces collocations dans la même forme, il faut en créer plusieurs.

Affiché (éd. papier et numérique) : « *partir per / partir entà* »

```
<form subtype="word" type="lemma" xml:lang="oc-lengadoc-grclass/oc-provenc-grclass" xml:id="form-125" n="1" cert="current" value="partir per">
```

```

<orth norm="partir" value="partir">partir</orth>
<colloc type="pre" norm=" per" value="per"> per</colloc>
...
</form>
<form subtype="word" type="lemma" xml:lang="oc-lengadoc-grclass/oc-provenc-grclass" xml:id="form-125" n="1" cert="current" value="partir per">
<orth norm="partir" value="partir">partir</orth>
<colloc type="pre" norm=" entà" value="entà"> entà</colloc>
...
</form>

```

3.2.5 Les formes fléchies et élidées

Elles se trouvent entre les balises **<form type="inflected">** et **</form>** ou **<form type="elided">** et **</form>** ou **<form type="inflectelided">** et **</form>**, à l'intérieur des balises de la forme principale.

Exemples :

Original : « beau, belle »

Affiché (éd. papier et numérique) : « beau, belle »

```

<form subtype="word" type="lemma" xml:lang="fr" xml:id="form-124" n="1" cert="ok" value="beau">
<orth norm="beau" value="beau" orig="beau">beau</orth>
...
<form subtype="word" type="inflected" xml:lang="fr" xml:id="form-125" n="1" cert="ok" value="belle">
<orth norm="belle" value="belle" orig="belle">belle</orth>
...
</form>
</form>

```

Original : « ornar (òrni) »

Affiché (éd. papier et numérique) : « ornar (òrni) »

```

<form subtype="word" type="lemma" xml:lang="oc-lengadoc-grclass" xml:id="form-1128" n="1" cert="current" value="ornar">
<orth norm="ornar" value="ornar" orig="ornar">ornar</orth>
<form subtype="word" type="inflected" xml:lang="oc-lengadoc-grclass" xml:id="form-1129" n="1" cert="current" value="òrni">
<orth norm="òrni" value="òrni" orig="òrni">òrni</orth>
<gramGrp>
<per norm="1"></per>
<number norm="singular"></number>
<tns norm="present"></tns>
<mood norm="indicative"></mood>
<gram type="catDet" norm="VerbelndPres1s"></gram>
<gram type="catGen" norm="Verbe"></gram>
</gramGrp>
...
</form>
...
</form>

```

Les dictionnaires considèrent parfois certains noms féminins comme une forme fléchie d'un nom masculin. Ce choix a été respecté lors de l'étiquetage de ces ouvrages.

Exemple :

Original : « charpentier, charpentière »

Affiché (éd. papier et numérique) : « charpentier, charpentière »

```
<form subtype="word" type="lemma" xml:lang="fr" xml:id="form-5423" n="1" cert="ok" value="charpentier">
  <orth norm="charpentier" value="charpentier" orig="charpentier">charpentier</orth>
  ...
  <form subtype="word" type="inflected" xml:lang="fr" xml:id="form-5424" n="1" cert="ok"
  value="charpentière">
    <orth norm="charpentière" value="charpentière" orig="charpentière">charpentière</orth>
    ...
  </form>
</form>
```

Les formes fléchies ne peuvent pas se trouver à l'intérieur d'autres formes fléchies ou de formes élidées. Les élisions de formes fléchies sont indiquées au même niveau que les formes fléchies ou élidées et elle prennent la valeur « **inflectelided** » à leur attribut « **type** ».

Exemple :

Original : « le, la, l' »

Affiché (éd. papier et numérique) : « le / l', la / l" »

```
<form subtype="word" type="lemma" xml:lang="fr" xml:id="form-30" n="1" cert="ok" value="le">
  <orth norm="le" value="le" orig="le">le</orth>
  <gramGrp>
    <gram type="catDet" norm="ArtDefMS"></gram>
    <gram type="catGen" norm="ArtDef"></gram>
  </gramGrp>
<form subtype="word" type="elided" xml:lang="fr" xml:id="form-31" n="1" cert="ok" value="l'">
  <orth norm="l'" value="l'" orig="l'">l'</orth>
  <gramGrp>
    <gram type="catDet" norm="ArtDefMS"></gram>
    <gram type="catGen" norm="ArtDef"></gram>
  </gramGrp>
</form>
<form subtype="word" type="inflected" xml:lang="fr" xml:id="form-32" n="1" cert="ok" value="la">
  <orth norm="la" value="la" orig="la">la</orth>
  <gramGrp>
    <gram type="catDet" norm="ArtDefFS"></gram>
    <gram type="catGen" norm="ArtDef"></gram>
  </gramGrp>
</form>
<form subtype="word" type="inflectelided" xml:lang="fr" xml:id="form-32" n="1" cert="ok" value="l"''>
  <orth norm="l"'' value="l"'' orig="l"''>l'</orth>
  <gramGrp>
    <gram type="catDet" norm="ArtDefFS"></gram>
    <gram type="catGen" norm="ArtDef"></gram>
  </gramGrp>
</form>
</form>
```

La flexion entre balises est indiquée telle qu'elle doit être affichée en sortie pour une édition papier. Le contenu de l'attribut « **expand** » indique comment la forme doit être affichée en sortie pour une édition numérique, si cet affichage est différent de celui de l'édition papier. Le contenu de l'attribut « **orig** » indique comment la forme est donnée dans le dictionnaire papier

d'origine (souvent, il donne uniquement la terminaison) ; l'attribut « **norm** » est utilisé pour indiquer la forme fléchie réelle (cf. 4).

Exemples :

Original : « *nuptial, -e (adj)* »

Affiché (éd. papier) : « *nuptial, -le (adj.)* »

Affiché (éd. numérique) : « *nuptial, nuptiale (adj.)* »

```
<form subtype="word" type="lemma" xml:lang="fr" xml:id="form-654" n="1" cert="current" value="nuptial">
  <orth norm="nuptial" value="nuptial" orig="nuptial">nuptial</orth>
  <form subtype="word" type="inflected" xml:lang="fr" xml:id="form-655" n="1" cert="current" value="nuptiale">
    <orth norm="nuptiale" value="nuptiale" orig="-e" expand="nuptiale">-le</orth>
    <gramGrp>
      <gen norm="feminine"></gen>
      <gram type="catDet" norm="AdjFS"></gram>
      <gram type="catGen" norm="Adj"></gram>
    </gramGrp>
  </form>
  <gramGrp>
    <pos norm="adjective" orig="adj"></pos>
    <gram type="catDet" norm="AdjMS"></gram>
    <gram type="catGen" norm="Adj"></gram>
    <lbl type="catGen" norm="adjectif" orig="aj" xml:lang="fr">adj.</lbl>
  </gramGrp>
  ...
</form>
```

Original : « *qualqu'un, -a (plur. : quelques uns, -unas)* »

Affiché (éd. papier et numérique) : « *qualqu'un, qualqu'una (f.), quelques-uns (pl.), quelques-unas (f. pl.)* »

```
<form subtype="word" type="lemma" xml:lang="oc-lengadoc-grclass" xml:id="form-3" n="1" cert="current" value="qualqu'un">
  <orth norm="qualqu'un" value="qualqu'un" orig="qualqu'un">qualqu'un</orth>
  <form subtype="word" type="inflected" xml:lang="oc-lengadoc-grclass" xml:id="form-4" n="1" cert="current" value="qualqu'una">
    <orth norm="qualqu'una" value="qualqu'una" orig="-a">qualqu'una</orth>
    <gramGrp>
      <gen norm="feminine" orig="fem."></gen>
      <gram type="catDet" norm="PronIndefFS"></gram>
      <gram type="catGen" norm="PronIndef"></gram>
      <lbl type="catDet" norm="féminin singulier" xml:lang="fr">f.</lbl>
    </gramGrp>
    ...
  </form>
  <form subtype="word" type="inflected" xml:lang="oc-lengadoc-grclass" xml:id="form-5" n="1" cert="current" value="quelques uns">
    <orth norm="quelques uns" value="quelques uns" orig="quelques uns">quelques uns</orth>
    <gramGrp>
      <number norm="plural" orig="plur."></number>
      <gram type="catDet" norm="PronIndefMP"></gram>
      <gram type="catGen" norm="PronIndef"></gram>
      <lbl type="catDet" norm="masculin pluriel" orig="plur" xml:lang="fr">pl.</lbl>
    </gramGrp>
```

```

...
</form>
<form subtype="word" type="inflected" xml:lang="oc-lengadoc-grclass" xml:id="form-6" n="1"
cert="current" value="quelques unas">
  <orth norm="quelques unas" value="quelques unas" orig="-unas">quelques unas</orth>
  <gramGrp>
    <gen norm="feminin"></gen>
    <number norm="plural"></number>
    <gram type="catDet" norm="PronIndefFP"></gram>
    <gram type="catGen" norm="PronIndef"></gram>
    <lbl type="catDet" norm="féminin pluriel" xml:lang="fr">f. pl.</lbl>
  </gramGrp>
  ...
</form>
...
</form>

```

3.2.6 Les informations grammaticales

On les indique entre les balises **<gramGrp>** et **</gramGrp>**. Elles contiennent les balises suivantes :

- **<case>** : cas, pour les pronoms personnels (voir en 6.1)
- **<gen>** : genre (masculine, feminine, neuter)
- **<gram>** : utilisé avec l'attribut **<gram type="polarity">** pour indiquer une forme affirmative (« **affirmative** ») ou négative (« **negative** ») d'un impératif ; sinon, utilisée pour indiquer la catégorie grammaticale de manière normée (voir en 3.2.6.1).
- **<iType>** : utilisé pour indiquer l'alternance vocalique des verbes, les mots invariables, le fait que le verbe soit ou pas inchoatif et les modèles de conjugaison (voir 3.2.6.3).
- **<mood>** : mode d'un verbe (**indicative**, **subjunctive**, **conditional**, **participle**, **impérative...**).
- **<note>** : informations diverses.
- **<number>** : nombre (**singular**, **plural**)
- **<per>** : personne (**1**, **2**, **3**)
- **<pos>** : « **part of speech** », catégorie grammaticale (**noun**, **adjective**, **adverb**, **pronoun...**)
- **<subc>** : sous-catégorie (pour les verbes : **transitive**, **intransitive...** ; pour les pronoms : **possessive**, **démonstrative...**)
- **<tns>** : temps d'un verbe (**past**, **present**, **future...**)

Les balises ci-dessus (sauf la balise **<note>**) ont un attribut « **norm** » qui contient leur valeur normalisée (cf. 4). Ils ne contiennent pas de texte brut. L'affichage de ces informations se fait dans le label de la catégorie grammaticale (voir en 3.2.6.2).

En général, les balises **<pos>**, **<gen>**, **<number>**, **<per>**, **<tns>**, **<mood>**, **<gram type="polarity">** ne sont pas utilisées, car elles sont redondantes avec les **<gram type="catDet">**. Elles sont parfois générées automatiquement à partir de la catégorie détaillée pour permettre le filtre par genre, par personne, par temps... par exemple pour les lexiques de conjugaisons.

Exemples :

<i>Nom commun féminin pluriel</i>

<gramGrp> <pos norm="noun"></pos>

```

<subc norm="common"></subc>
<gen norm="feminine"></gen>
<number norm="plural" orig="pluriel"></number>
</gramGrp>

```

« *abaishem* » (1e pers. pl. imp. pst pos.)

```

<form subtype="word" type="inflected" xml:lang="oc-gascon-grclass" xml:id="form-1" n="1" cert="ok"
value="abaishem">
<orth norm="abaishem" value="abaishem" orig="abaishem">abaishem</orth>
<gramGrp>
<per norm="1"></per>
<number norm="plural"></number>
<mood norm="imperative"></mood>
<tns norm="present"></tns>
<gram type="polarity" norm="affirmative"></gram>
</gramGrp>
...
</form>

```

3.2.6.1 Les catégories grammaticales normalisées

La sous-balise **<gram>** est utilisée pour indiquer de manière normalisée la catégorie grammaticale de la forme. Elle est complétée par un attribut « **type** » qui indique la norme utilisée (et le degré de précision de l'information) :

- **catGen** : catégorie grammaticale générale (voir en 5.1). Il s'agit de la PoS, la catégorie que partagent le lemme et ses formes fléchies.
- **catDet** : catégorie grammaticale détaillée (voir en 5.1). Il s'agit de la PoS et des informations morphosyntaxiques, qui dépendent du contexte (genre pour les adjectifs, nombre pour les noms, personne, temps et mode pour les verbes...)
- **catFam** : famille grammaticale (voir en 5.1). Utilisée notamment dans les lexiques destinés au machine learning. Grande famille grammaticale à laquelle appartient le mot (ex. : adjectif, déterminant, pronom...).
- **catEagles** : catégorie exprimée en fonction de la norme *Eagles* (voir en 5.1.13), utilisée uniquement dans certains lexiques spécifiques.
- **catGrace** : catégorie exprimée en fonction de la norme *Grace* (voir Vergez-Couret, 2016, *Description du lexique Lofloc*), utilisée uniquement dans certains lexiques spécifiques.

Les ressources du Congrès utilisent systématiquement les catégories détaillées et les catégories générales. Celles-ci peuvent être complétées par les codes Grace et Eagles utilisés par les partenaires.

On n'a qu'un **<gram type="catDet">**, qu'un **<gram type="catGen">** et qu'un **<gram type="catFam">** par **<gramGrp>**. Par contre, on peut avoir plusieurs **<gram type="catEagles">** et plusieurs **<gram type="catGrace">**.

Exemple :

Nom commun masculin singulier

```

<gramGrp>
<gram type="catEagles" norm="N11100"></gram>
<gram type="catGrace" norm="Ncms"></gram>

```

```

<gram type="catGen" norm="NomComM"></gram>
<gram type="catDet" norm="NomComMS"></gram>
<gram type="catFam" norm="NomCom"></gram>
</gramGrp>

```

3.2.6.2 L'affichage des catégories grammaticales

Pour gérer l'affichage des catégories grammaticales, on utilise la balise **<lbl>**. Elle comporte un attribut « **type** » qui indique si l'affichage concerne la catégorie grammaticale générale (**catGen**) ou la catégorie grammaticale détaillée (**catDet**).

Elle comporte un attribut « **norm** » qui donne la forme pleine que l'on veut afficher pour la catégorie (ex. **nom féminin**). S'il n'y a rien à afficher pour la catégorie, il n'y a pas d'attribut « **norm** ». Le contenu entre balises, lui, contient la forme abrégée (ex. : **nf**) à afficher pour une édition papier. Si l'affichage pour une édition numérique est différent de celui pour une édition papier, on utilise l'attribut « **expand** ». L'affichage de la catégorie tel que dans le dictionnaire original est indiqué dans l'attribut « **orig** ».

Enfin, la balise comporte un attribut « **xml:lang** » qui indique la langue du contenu à afficher.

Exemple :

Nom commun masculin singulier

```

<gramGrp>
  <lbl type="catGen" norm="nom comun feminin" xml:lang="oc-provenc-grclass" expand="n.
fem.">nf.</lbl>
  <lbl type="catGen" norm="nom commun féminin" xml:lang="fr" expand="n. fém.">nf.</lbl>
  <lbl type="catDet" norm="singular" xml:lang="oc-provenc-grclass"></lbl>
  <lbl type="catDet" norm="singulier" xml:lang="fr"></lbl>
</gramGrp>

```

Si l'affichage concerne plusieurs langues et/ou plusieurs variétés, on les indique toutes dans l'attribut « **xml:lang** » en les séparant par « / ».

Exemple :

Nom commun masculin singulier

```

<gramGrp>
  <lbl type="catGen" norm="nom comun masculin" xml:lang="fr/oc-lengadoc-grclass/oc-gascon-grclass"
expand="n. masc.">nm.</lbl>
  <lbl type="catDet" norm="singulier" xml:lang="fr/oc-lengadoc-grclass/oc-gascon-grclass"></lbl>
</gramGrp>

```

3.2.6.3 La balise **<iType>**

La balise **<iType>**, qui se trouve dans la balise **<gramGrp>**, a été utilisée pour encadrer le type d'alternance des verbes. Elle a dans ce cas été taguée **<iType type="altvb">**. Elle a également été utilisée pour indiquer qu'un mot est invariable. Elle est alors taguée **<iType type="accord">**. Pour les tableaux de modèles de conjugaison de verbes ou de déclinaisons, elle est taguée **<iType type="paradigm">**. Pour l'« inchoativité », elle est taguée **<iType type="inchoative">**.

Le contenu des balises **<iType>** est normalisé. On utilise l'attribut « **norm** » si le contenu entre les balises ne correspond pas aux normes définies (voir 6.4 et 6.5). On utilise l'attribut « **orig** »

si le contenu dans le dictionnaire d'origine ne correspond pas à ce qu'on veut afficher (cf. 4). On utilisera l'attribut « **expand** » si on veut indiquer que dans une édition numérique éventuelle, le contenu à afficher sera différent de celui destiné à une édition papier (qui se trouve entre les balises).

Pour la balise **<iType type="paradigm">**, on peut ajouter un attribut « **subtype** » qui permet de préciser la nature ou la source du paradigme.

Exemples :

Affiché (éd. papier) : « *fondar* (o) »

Affiché (éd. numérique) : « *fondar* (o/o) »

```
<form subtype="word" type="lemma" xml:lang="oc-lengadoc-grclass" xml:id="form-354" n="1"
cert="current" value="fondar">
<orth norm="fondar" value="fondar">fondar</orth>
<gramGrp>
  <iType type="altvb" norm="o/o" expand="o/o">o</iType>
</gramGrp>
...
</form>
```

Affiché (éd. papier et numérique) : « *carmin* (adj. inv.) »

```
<form subtype="word" type="lemma" xml:lang="fr" xml:id="form-28" n="1" cert="current" value="camin">
<orth norm="carmin" value="carmin">carmin</orth>
<gramGrp>
  <lbl type="catGen" norm="adjective" xml:lang="fr">adj.</lbl>
  <iType type="accord" norm="invariable">inv.</iType>
</gramGrp>
...
</form>
```

Affiché (éd. papier et numérique) : « *manjar* (103) »

```
<form subtype="word" type="lemma" xml:lang="oc-lengadoc-grclass" xml:id="form-571" n="1"
cert="current" value="manjar">
<orth norm="manjar" value="manjar">manjar</orth>
<gramGrp>
  <iType type="paradigm" subtype="vèrb'Òc" norm="lengadocien 103">103</iType>
</gramGrp>
...
</form>
```

Affiché (éd. papier et numérique) : « *apelar* (v., e/è) »

```
<form subtype="word" type="lemma" xml:lang="oc-lengadoc-grclass" xml:id="form-5874" n="1"
cert="current" value="apelar">
<orth norm="apelar" value="apelar">apelar</orth>
<gramGrp>
  <lbl type="catGen" norm="verbe" xml:lang="fr">v.</lbl>
</gramGrp>
<gramGrp>
  <iType type="altvb" norm="e/è">e/è</iType>
```

```
</gramGrp>
...
</form>
```

Affiché (éd. papier et numérique) : « *chatir* (-iss-) »

```
<form subtype="word" type="lemma" xml:lang="oc-auvern-grclass" xml:id="form-8" n="1" cert="current"
value="chatir">
<orth norm="chatir" value="chatir" >chatir</orth>
<gramGrp>
  <iType type="inchoative" norm="inchoative">-iss-</iType>
</gramGrp>
...
</form>
```

Si la balise **<iType>** est contenue dans une forme multivariétés mais ne concerne pas toutes les variétés de la forme, on peut utiliser l'attribut « **xml:lang** » avec les valeurs des codes langues de l'IANA (voir en 7.1) pour indiquer la ou les variétés auxquelles s'applique l'information. Si plusieurs variétés ont la même information **<iType>**, on sépare les langues par un « / ».

Si toutes les variétés de la forme ont la même information **<iType>**, en revanche, pas besoin de répéter la balise. Il suffit de ne pas mettre l'attribut « **xml:lang** ».

Affiché (éd. papier et numérique) : « *faiçonar* (o/o) ; *formar* (gasc. : ò/o, leng. : ò/o, prov. : o/o) »

```
<cit type="translation">
<form subtype="word" type="lemma" xml:lang="oc-gascon-grclass/oc-lengadoc-grclass/oc-provenc-
grclass" xml:id="form-2" n="1" cert="current" value="faiçonar">
<orth norm="faiçonar" value="faiçonar" >faiçonar</orth>
<gramGrp>
  <iType type="altvb" norm="o/o">o/o</iType>
</gramGrp>
</form>
</cit>
<cit type="translation">
<form subtype="word" type="lemma" xml:lang="oc-gascon-grclass/oc-lengadoc-grclass/oc-provenc-
grclass" xml:id="form-3" n="1" cert="current" value="formar">
<orth norm="formar" value="formar" >formar</orth>
<gramGrp>
  <iType type="altvb" norm="ò/o" xml:lang="oc-gascon-grclass/oc-lengadoc-grclass">ò/o</iType>
  <iType type="altvb" norm="o/o" xml:lang="oc-provenc-grclass">o/o</iType>
</gramGrp>
</form>
</cit>
```

3.2.7 La ou les prononciations

Elles sont placées entre les balises **<pron>** et **</pron>**. Si celle du dictionnaire source ne l'est pas, une prononciation normalisée peut être indiquée grâce à l'attribut « **norm** ».

La balise est complétée par l'attribut « **extent** » qui prend les valeurs suivantes :

- **full** : prononciation du mot complet
- **part** : prononciation d'une partie du mot

- pref : prononciation du préfixe
- suf : prononciation du suffixe
- poorRhyme : rime pauvre
- richRhyme : rime riche

La notation utilisée pour écrire prononciation (IPA, phonétique...) peut être indiquée grâce à l'attribut « **notation** ».

La balise **<pron>** porte les attributs « **norm** », « **orig** » et « **expand** » (cf. 4).

Exemples :

Affiché (éd. papier et numérique) : « agòra [a'γɔra] »

```
<form subtype="word" type="lemma" xml:lang="oc-gascon-grclass" xml:id="form-853" n="1" cert="ok"
value="agòra">
  <orth norm="agòra" value="agòra">agòra</orth>
  <pron notation="IPA" norm="a'γɔra">a'γɔra</pron>
  ...
</form>
```

Original : « coquilhar [-ar] »

Affiché (éd. papier) : « coquilhar [-ar] »

Affiché (éd. numérique) : « coquilhar [-'jar] »

```
<form subtype="word" type="lemma" xml:lang="oc-gascon-grclass" xml:id="form-183" n="1" cert="ok"
value="coquilhar">
  <orth norm="coquilhar" value="coquilhar" orig="coquilhar">coquilhar</orth>
  <pron extent="part" norm="-'ar" orig="-'ar" expand="-'jar">'-ar</pron>
  ...
</form>
```

Original : « jornal [journal] »

Affiché (éd. papier et numérique) : « jornal [dʒurnal] »

```
<form subtype="word" type="lemma" xml:lang="oc-gascon-grclass" xml:id="form-456" n="1" cert="ok"
value="jornal">
  <orth norm="jornal" value="jornal" orig="jornal">jornal</orth>
  <pron norm="dʒurnal" orig="jornal">dʒurnal</pron>
  ...
</form>
```

Si la forme de la prononciation est multivariété, mais que la prononciation ne concerne qu'une seule des variétés, on peut l'indiquer grâce à l'attribut « **xml:lang** ».

Si la forme se prononce de la même manière dans plusieurs variétés mais pas toutes, on les indique toutes dans l'attribut « **xml:lang** » en les séparant par « / ».

Original : « harèm [harɛm] [arem] »

Affiché (éd. papier et numérique) : « harèm [harɛm] [arem] »

```
<form subtype="word" type="lemma" xml:lang="oc-gascon-grclass /oc-lengadoc-grclass/oc-provenc-
grclass" n="1" xml:id="form-876" cert="ok" value="harèm">
```

```

<orth norm="harèm" value="harèm">harèm</orth>
<pron norm="harèm" orig="harem" xml:lang="oc-gascon">harem</pron>
<pron norm="arèm" orig="arem" xml:lang="oc-lengadoc/oc-vivaraup">arèm</pron>
</form>

```

3.2.8 La syllabisation

On peut indiquer l'orthographe du mot avec son découpage en syllabe grâce aux balises **<syll>** et **</syll>** (les syllabes sont alors découpées avec des !).

La balise est complétée par l'attribut « **extent** » qui prend les valeurs suivantes :

- **full** : syllabisation du mot complet
- **part** : syllabisation d'une partie du mot
- **pref** : syllabisation du préfixe
- **suff** : syllabisation du suffixe

La balise **<syll>** porte les attributs « **norm** », « **orig** » et « **expand** » (cf. 4).

Exemples :

Affiché (éd. papier et numérique) : « *area (re|a)* »

```

<form subtype="word" type="lemma" xml:lang="oc-gascon-grclass" xml:id="form-25" n="1" cert="ok"
value="area">
  <orth norm="area" value="area">area</orth>
  <syll extent="part" norm="re|a">re|a</pron>
  ...
</form>

```

Original : « *aviar (a|viar)* »

Affiché (éd. papier et numérique) : « *aviar (a/viar)* »

```

<form subtype="word" type="lemma" xml:lang="oc-gascon-grclass" xml:id="form-22" n="1" cert="ok"
value="aviar">
  <orth norm="aviar" value="aviar" orig="aviar">aviar</orth>
  <syll orig="a|viar" norm="a|viar">a|viar</pron>
  ...
</form>

```

3.2.9 Les informations d'usage

Pour indiquer dans quel contexte un mot est utilisé, on se sert des balises **<usg>** et **</usg>**. Elles sont obligatoirement complétées par l'attribut « **type** » qui peut prendre les valeurs suivantes :

- **ant** : antonyme
- **binName** : dénomination binomiale (nom de genre + nom d'espèce) pour les plantes et les animaux
- **colloc** : collocation usuelle
- **dom** : domaine ou sujet (géographie, théâtre, sport...)
- **geo** : aire géographique
- **gram** : usage grammatical (ex : après un nom, suivi d'un subjonctif...)
- **hint** : information non classifiable

- **hyper** : hyperonyme, catégorie à laquelle appartient le mot (ex : « animal » pour « chat »)
- **lang** : langue, la plupart du temps langue d'origine (francisme, catalanisme), indiquée par un code ISO dans l'attribut « **norm** ».
- **plev** : niveau d'usage (principalement, habituellement...) ou de sens (général, principal...)
- **register** : registre de langue (familier, trivial, littéraire...)
- **style** : style (sens figuré, littéral...)
- **synonym** : synonyme
- **time** : époque où était/est utilisé le terme (archaïque, vieux, 15e s., néologisme...)

Les informations d'usage (zones géographiques, langues, domaines...) sont normalisées grâce à l'attribut « **norm** » (cf. 6.15 pour « **dom** », 6.16 pour « **plev** », 6.17 pour « **register** », 6.18 pour « **style** », 6.19 pour « **time** »).

La balise **<usg>** accepte également les attributs « **orig** » et « **expand** » (cf. 4).

Exemples :

<i>Original : « amocher (fam., fig.) »</i>
<i>Affiché (éd. papier) : « amocher (fam., sens fig.) »</i>
<i>Affiché (éd. numérique) : « amocher (familier, sens figuré) »</i>
<pre><form subtype="word" type="lemma" xml:lang="fr" xml:id="form-756" n="1" cert="ok" value="amocher"> <orth norm="amocher" value="amocher" orig="amocher">amocher</orth> <usg type="register" norm="familiar" orig="fam." expand="familier">fam.</usg> ... </form> <sense xml:id="sense-623" level="1"> <usg type="style" norm="figurative" orig="fig." expand="sens figuré">sens fig.</usg> </sense></pre>

<i>Original : « pinganada {V-B} »</i>
<i>Affiché (éd. papier et numérique) : « pinganada (Vic-Bilh) »</i>
<pre><form subtype="word" type="lemma" xml:lang="oc-gascon-grclass" xml:id="form-236" n="1" cert="ok" value="pinganada"> <orth norm="pinganada" value="pinganada" orig="pinganada">pinganada</orth> <usg type="geo" norm="Vic-Bilh" orig="V-B">Vic-Bilh</usg> ... </form></pre>

Si l'information d'usage ne concerne que certaines variétés d'une forme qui existe dans plusieurs variétés, on peut l'indiquer grâce à l'attribut « **xml:lang** ». Il prend en valeur un ou plusieurs codes langues de l'IANA (voir en 7.1) séparés par « / ».

<i>Affiché (éd. papier et numérique) : « chepic (rare en languedocien) »</i>
<pre><form subtype="word" type="lemma" xml:lang="oc-gascon-grclass" xml:id="form-658" n="1" cert="ok" value="chepic"> <orth norm="chepic" value="chepic">chepic</orth> <usg type="plev" norm="rare" xml:lang="oc-lengadoc">rare en languedocien</usg> ... </form></pre>

La balise **<usg>** a été, en général, placée sous les balises **<form> ... </form>** lorsqu'elle indiquait la zone géographique, la langue d'origine, le registre, le niveau d'usage, l'ère temporelle, tout ce qui concerne la forme elle-même, indépendamment de son contexte d'utilisation. Elle a été placée sous les balises **<sense> ... </sense>** ou **<cit> ... </cit>** lorsqu'elle indiquait le reste (domaine, style, synonyme, hyperonyme, registre lié au sens...), ce qui concerne le sens.

Exemples :

Affiché (éd. papier et numérique) : « *merde (fam)* »

```
<form subtype="word" type="lemma" xml:lang="fr" xml:id="form-5312" n="1" cert="current" value="merde">
  <orth norm="merde" value="merde">merde</orth>
  <usg type="register" norm="familiar">fam</usg>
  ...
</form>
```

Original : « *épater (fig/fam)* »

Affiché (éd. papier et numérique) : « *épater (fig., fam.)* »

```
<form subtype="word" type="lemma" xml:lang="fr" xml:id="form-8" n="1" cert="current" value="épater">
  <orth norm="épater" value="épater" orig="épater">épater</orth>
  ...
</form>
<sense xml:id="sense-1" level="1">
  <usg type="style" norm="litteral" orig="prop">prop.</usg>
  ...
</sense>
<sense xml:id="sense-5" level="1">
  <usg type="style" norm="figurative" orig="fig">fig.</usg>
  <usg type="register" norm="familiar" orig="fam">fam.</usg>
  ...
</sense>
```

Affiché (éd. papier et numérique) : « *suturer (couture)* »

```
<form subtype="word" type="lemma" xml:lang="fr" xml:id="form-564" n="1" cert="current" value="suturer">
  <orth norm="suturer" value="suturer">suturer</orth>
  ...
</form>
<sense xml:id="sense-520" level="1">
  <usg type="dom" norm="couture">couture</usg>
  ...
</sense>
```

Affiché (éd. papier et numérique) : « *baiser (fam.)* »

```
<form subtype="word" type="lemma" xml:lang="fr" xml:id="form-12" n="1" cert="current" value="baiser">
  <orth norm="baiser" value="baiser">baiser</orth>
  ...
</form>
<sense xml:id="sense-1" level="1">
  <usg type="plev" norm="usually">gén.</usg>
  ...
</sense>
```

```

<sense xml:id="sense-10" level="1">
  <usg type="register" norm="familiar">fam.</usg>
  ...
</sense>

```

3.2.10 Les informations de localisation

Lorsqu'une forme représente un lieu, on peut y adjoindre une ou plusieurs localisations entre les balises **<location>** et **</location>**.

Elles peuvent contenir une indication de pays entre les balises **<country>** et **</country>** (voir en 3.2.3).

Elles peuvent contenir une indication de subdivision territoriale plus précise entre les balises **<region>** et **</region>** (voir en 3.2.3).

Elles peuvent contenir une indication de code commune dans l'attribut « **key** » d'une **<settlement type="city"></settlement>**.

Elles peuvent contenir une indication de code postal entre les balises **<address><postCode>** et **</postCode></address>**.

Pour indiquer les valeurs de **<country>** et **<region>**, on utilise les balises **<reg>**, **<orig>** et **<expan>** (voir 4).

La balise **<location>** peut également contenir des coordonnées géographiques entre les balises **<geo>** et **</geo>**. On y insère une latitude et une longitude séparées par un espace.

Exemple :

Affiché (éd. papier et numérique) : « Pau (Pyrénées-Atlantiques, France) {43.3219001 -0.3786763} »

```

<form subtype="word" type="lemma" xml:lang="oc-gascon-grclass" xml:id="form-521" n="1" cert="ok"
value="Pau">
  <orth norm="Pau" value="Pau">Pau</orth>
  <location>
    <region type="department" key="64">
      <choice>
        <reg>Pyrénées-Atlantiques</reg>
      </choice>
    </region>
    <country key="FR">
      <choice>
        <reg>France</reg>
      </choice>
    </country>
    <settlement type="city" key="64445"></settlement>
    <address>
      <postCode>64000</postCode>
    </address>
    <geo>43.3219001 -0.3786763</geo>
  </location>
  ...
</form>

```

3.2.11 Informations diverses

Elles sont insérées entre les balises **<note>** et **</note>**. Ces balises ne se trouvent qu'à l'intérieur de balises ne pouvant contenir ni du texte brut, ni les balises **<reg>**, **<orig>** et **<expan>** (voir 4).

La balise **<note>** est complétée par un attribut « **rend** » qui indique si on en affiche ou pas (**visible**, **invisible**) le texte dans la version finale du dictionnaire. Une note « **invisible** » sera affichée lorsqu'on essaiera de reconstituer la version d'origine, mais elle sera masquée si on affiche une version harmonisée du dictionnaire dans laquelle l'esthétique prime sur la fidélité à la source.

Elle comporte également un attribut « **place** » qui indique si elle est située avant (« **above** ») ou après (« **below** ») le reste du contenu de la balise parente (si la note est le seul enfant de la balise parente, on met la valeur « **above** »).

Exemples :

Original : « Voir aussi : batailleur, bagarreur »

Affiché (éd. papier et numérique) : « cf. : batailleur, bagarreur »

```
<xr type="cf">
<note rend="invisible" place="above">Voir aussi :</note>
<lbl type="reference" norm="cf">Cf.</lbl>
<cit type="reference">
  <form subtype="word" type="lemma" xml:lang="fr" cert="current" xml:id="form-350" n="1" cert="ok"
value="batailleur">
    <orth norm="batailleur" value="batailleur" orig="batailleur">batailleur</orth>
    <gramGrp>
      <gram type="catDet" norm="AdjMS"></gram>
      <gram type="catGen" norm="Adj"></gram>
    </gramGrp>
  </form>
</cit>
<cit type="reference">
  <form subtype="word" type="lemma" xml:lang="fr" cert="current" xml:id="form-351" n="1" cert="ok"
value="bagarreur">
    <orth norm="bagarreur" value="bagarreur" orig="bagarreur">bagarreur</orth>
    <gramGrp>
      <gram type="catDet" norm="AdjMS"></gram>
      <gram type="catGen" norm="Adj"></gram>
    </gramGrp>
  </form>
</cit>
</xr>
```

Affiché (éd. papier et numérique) : « Le pronom sujet n'est pas traduit en occitan »

```
<cit type="translation">
<note rend="visible" place="above">Le pronom sujet n'est pas traduit en occitan</orth>
</form>
```

Original : « grelhon ; on peut aussi utiliser : nina (litt., pupille) »

Affiché (éd. papier et numérique) : « *grelhon* ; *nina* (*litt., pupille*) »

```
<cit type="translation">
  <form subtype="word" type="lemma" xml:lang="oc-gascon-grclass" xml:id="form-24" n="1" cert="ok"
  value="grelhon">
    <orth norm="grelhon" value="grelhon" orig="grelhon">grelhon</orth>
    ...
  </form>
</cit>
<note rend="invisible" place="above">on peut aussi utiliser</note>
<cit type="translation">
  <form subtype="word" type="lemma" xml:lang="oc-gascon-grclass" xml:id="form-25" n="1" cert="ok"
  value="nina">
    <orth norm="nina" value="nina" orig="nina">nina</orth>
    ...
  </form>
<note rend="visible" place="below">litt., pupille</note>
</cit>
```

Original : « (ex : *c'est drôlement bon*) »

Affiché (éd. papier et numérique) : « *c'est drôlement bon* »

```
<re xml:id="re-587" type="example">
  <note rend="invisible" place="above">ex</note>
  <form subtype="phrase" type="lemma" xml:lang="fr" xml:id="form-874" n="1" cert="current" value="c'est
  drôlement bon">
    <orth norm="c'est drôlement bon" value="c'est drôlement bon" orig="c'est drôlement bon">c'est
  drôlement bon</orth>
  ...
</form>
</re>
```

Original : « *m' : me apostrophé* »

Affichage : « *m' : me apostrophé* »

```
<form subtype="word" type="lemma" xml:lang="fr" xml:id="form-736" n="1" cert="current" value="me">
  <form subtype="word" type="elided" xml:lang="fr" xml:id="form-737" n="1" cert="current" value="m">
    <orth norm="m'" value="m'" orig="m'">m'</orth>
    <note rend="visible" place="below">me apostrophé</note>
    ...
  </form>
  <orth norm="me" value="me"></orth>
  ...
</form>
```

Dans de très rares cas (exceptions « en vrac » des dictionnaires historiques notamment), le contenu des balises **<note>** peut comporter des balises **<hi style="style css"></hi>** qui encadrent du texte ayant un style particulier. Mais si le document doit être utilisé ensuite par une machine, cela perturbe le fonctionnement.

Exemple :

« *Je suis vraiment pressé* »

```
<note rend="visible" place="below">Je suis <hi style="font-weight:bold; text-
```

decoration:underline;">>**vraiment**</hi> pressé</note>

3.3 La personne-vedette

Si l'entrée n'est pas à propos d'un mot mais d'une personne, on a à la place de la balise **<form>** une balise **<person>**.

La balise **<person>** est complétée par un attribut « **xml:id** » qui est un identifiant unique pour les personnes d'un document. Il prend la forme **person-[nombre]**. Elle comporte aussi un attribut « **sortkey** » qui indique la chaîne de caractère représentant la personne pour le tri alphabétique (souvent nom_prénom ou nom_prénom_particule).

Les noms, prénoms et autres dénominations de la personne sont indiquées à l'intérieur d'une balise **<persName>**. Celle-ci peut comporter un attribut « **xml:lang** » pour le cas où l'on voudrait indiquer plusieurs noms de la personne correspondant à plusieurs langues. La balise **<persname>** peut contenir les sous-balises suivantes :

- **<forename>** : prénom
- **<surname>** : nom de famille
- **<addName>** : nom additionnel, comme l'épithète « le Grand » pour « Alexandre le Grand », le titre honorifique « Mamie » pour « Mamie Eisenhower », un nom de scène, un surnom... On précise le type de nom additionnel (**honorific**, **epithet**, **nick**, **stagename**) dans un attribut « **type** ».
- **<roleName>** : pour les particules honorifiques ou titres (Mme, Dr, Colonel...)
- **<genName>** : pour les indications de génération (numéros des rois, « junior », « l'ancien », etc.)
- **<nameLink>** : les particules qui font le lien entre les noms (ex. : « de », « van der »...)
- **<choice>** : les sous-balises **<reg>**, **<orig>** et **<expan>** (voir 4) servent à indiquer d'un seul tenant le nom affiché et le nom original.

Toutes ces sous-balises, en-dehors de **<choice>**, ont un attribut « **sort** » qui permet d'indiquer l'ordre dans lequel il faut les assembler.

Original : « Ferry, Jules »

Affiché (éd. papier) : « J. Ferry »

Affiché (éd. numérique) : « Jules Ferry »

```
<person xml:id="person-12" sortkey="ferry_jules">
  <persName>
    <forename sort="1">Jules</forename>
    <surname sort="2">Ferry</surname>
    <choice>
      <reg>J. Ferry</reg>
      <orig>Ferry , Jules</orig>
      <expan>Ferry , Jules</expan>
    </choice>
  </persName>
</person>
```

Les dates et lieux de naissance et de mort sont indiqués grâce aux balises **<birth>** et **<death>**, qui peuvent contenir les sous-balises **<date>** (voir en 3.2.3) et/ou **<location>** (voir en 3.2.10).

Affiché (éd. papier et numérique) : « Frédéric Mistral (8 septembre 1830, Maillane - 25 mars 1914, Maillane) »

```
<person xml:id="person-635" sortkey="mistral_frédéric">
<persName>
  <forename sort="1">Frédéric</forename>
  <surname sort="2">Mistral</surname>
  <choice>
    <reg>Frédéric Mistral</reg>
  </choice>
</persName>
<birth>
  <date when="1830">
    <choice>
      <reg>8 septembre 1830</reg>
    <choice>
    </date>
<location>
  <settlement type="city" key="13052">
    <choice>
      <reg>Maillane</reg>
    <choice>
    </settlement>
</location>
</birth>
<death>
  <date when="1914">
    <choice>
      <reg>25 mars 1914</reg>
    <choice>
    </date>
<location>
  <settlement type="city" key="13052">
    <choice>
      <reg>Maillane</reg>
    <choice>
    </settlement>
</location>
</death>
</person>
```

Les informations liées aux occupations de la personne (emploi, activité associative ou artistique...) sont indiquées grâce à la balise **<occupation>**. Celle-ci peut comporter des attributs « **from** » et « **to** » qui indiquent la date de début et/ou la date de fin de la période pendant laquelle la personne a exercé l'occupation. Les sous-balises **<roleName>** et **<orgName>** indiquent respectivement le poste occupé et la structure dans les sous-balises **<reg>** et/ou **<orig>** et/ou **<expand>** d'une sous-balise **<choice>** (voir en 4).

Affiché (éd. papier et numérique) : « Pr. Gilbert Mercadier (président, Lo Congrès) »

```
<person xml:id="person-32" sortkey="mercadier_gilbert">
<persName>
  <roleName sort="1">professeur</roleName>
  <forename sort="2">Gilbert</forename>
  <surname sort="3">Mercadier</surname>
  <choice>
    <reg>Pr. Gilbert Mercadier</reg>
  </choice>
</persName>
```

```

<occupation from="2011">
  <roleName>
    <choice>
      <reg>Président</reg>
    </choice>
  </roleName>
  <orgName>
    <choice>
      <reg>Le Congrès</reg>
    </choice>
  </orgName>
</occupation>
</person>

```

Les informations sur le lieu de résidence d'une personne sont indiquées dans la balise **<residence>**. Celle-ci peut comporter des attributs « **from** » et « **to** » qui indiquent la date de début et/ou la date de fin de la période pendant laquelle la personne a résidé dans le lieu désigné. Le lieu est indiqué dans une sous-balise **<location>** (voir en 3.2.10).

Affiché (éd. papier et numérique) : « *Louis IX dit « Saint Louis » (Poissy)* »

```

<person xml:id="person-32" sortkey="louis_9">
  <persName>
    <forename sort="1">Louis</forename>
    <genName sort="2">IX</genName>
    <addName sort="3" type="nick">Saint Louis</addName>
    <choice>
      <reg>Louis IX dit « Saint Louis »</reg>
    </choice>
  </persName>
  <residence>
    <location>
      <settlement type="city">
        <choice>
          <reg>Poissy</reg>
        </choice>
      </settlement>
    </location>
  </residence>
</person>

```

Les informations sur le genre de la personne sont indiquées dans la balise **<gender>**, via les sous-balises **<reg>**, **<orig>** et **<expan>** d'une sous-balise **<choice>** (voir en 4). Celle-ci comporte un attribut « **value** » dans laquelle le genre est indiqué de manière normalisée (**masculine**, **feminine**, **other**) et éventuellement du texte brut où l'on a l'information de genre telle qu'elle sera affichée pour le public.

Affiché (éd. papier et numérique) : « *Hermaphrodite (dieu des deux sexes)* »

```

<person xml:id="person-1456" sortkey="hermaphrodite">
  <persName>
    <forename sort="1">Hermaphrodite</forename>
    <choice>
      <reg>Hermaphrodite</reg>
    </choice>
  </persName>
  <occupation>
    <roleName>

```

```

<choice>
  <reg>dieu</reg>
</choice>
</roleName>
</occupation>
<gender value="other">
  <choice>
    <reg>des deux sexes</reg>
  </choice>
</gender>
</person>

```

La nationalité de la personne est indiquée grâce à la balise **<nationality>**. Celle-ci comporte un attribut « **key** » qui contient un code ISO 3166 de pays. Elle contient une balise **<choice>** qui contient les sous-balises **<reg>** et/ou **<orig>** et/ou **<expand>**.

Affiché (éd. papier et numérique) : « *Mme de la Fayette (romancière française)* »

```

<person xml:id="person-14" sortkey="fayette_de_la">
  <persName>
    <addName sort="1" type="honorific">Mme</addName>
    <nameLink sort="2">de la</nameLink>
    <surname sort="3">Fayette</surname>
    <choice>
      <reg>Mme de la Fayette</reg>
    </choice>
  </persName>
  <occupation>
    <roleName>
      <choice>
        <reg>romancière</reg>
      </choice>
    </roleName>
  </occupation>
  <nationality key="FR">
    <choice>
      <reg>française</reg>
    </choice>
  </nationality>
</person>

```

Si l'on veut indiquer les langues connues par la personne, on peut les indiquer dans la balise **<langKnowledge>** qui contient elle-même :

- un **<lbl type="langKnowledge">** éventuel, pour l'affichage de l'information. Il accepte les attributs « **norm** », « **orig** » et « **expand** » (cf. 4).
- une ou plusieurs balises **<langKnown>**. Celle-ci comporte un attribut obligatoire « **tag** » avec le code langue de l'IANA de la langue connue, et un attribut facultatif « **level** » qui peut préciser le niveau de maîtrise de la langue. Elle contient une balise **<choice>** qui contient les sous-balises **<reg>** et/ou **<orig>** et/ou **<expand>**.

Affiché (éd. papier) : « *Danís Chapduelh / Denis Chadeuil (lang. : oc. lim., oc. lang., fr.)* »

Affiché (éd. numérique) : « *Danís Chapduelh / Denis Chadeuil (langues : occitan limousin, occitan languedocien, français)* »

```

<person xml:id="person-8" sortkey="chapduelh_danis">
  <persName xml:lang="oc">

```

```

<forename sort="1">Danís</forename>
<surname sort="2">Chapduelh</surname>
<choice>
  <reg>Danís Chapduelh</reg>
</choice>
</persName>
<persName xml:lang="oc">
  <forename sort="1">Denis</forename>
  <surname sort="2">Chadeuil</surname>
  <choice>
    <reg>Denis Chadeuil</reg>
  </choice>
</persName>
<langKnowledge>
  <lbl type="langKnowledge" expand="lang.">langues</lbl>
  <langKnown level="native" tag="oc-lemosin">
    <choice>
      <reg>oc. lim.</reg>
      <expan>occitan limousin</expan>
    </choice>
  </langKnown>
  <langKnown level="C1" tag="oc-lengadoc">
    <choice>
      <reg>oc. lang.</reg>
      <expan>occitan languedocien</expan>
    </choice>
  </langKnown>
  <langKnown level="native" tag="fr">
    <choice>
      <reg>fr.</reg>
      <expan>français</expan>
    </choice>
  </langKnown>
</langKnowledge>
</person>

```

On peut relier la personne à un identifiant dans un répertoire grâce à la balise **<idno>**. Celle-ci comporte un attribut « **type** » qui permet d'indiquer la nature de l'identifiant. On peut aussi lui lier un fichier ou une URL grâce à la balise **<ptr>** (voir en 3.6). Enfin, la balise **<person>** peut contenir des informations diverses dans une sous-balise **<note>**.

Affiché (éd. papier et numérique) : « Alexandre lo Grand plan coneget en euròpa »

```

<person xml:id="person-124" sortkey="grand_alexandre_le">
  <persName xml:lang="oc">
    <forename sort="1">Alexandre</forename>
    <addName sort="2" type="epithet">lo Grand</addName>
    <choice>
      <reg>Alexandre lo Grand</reg>
    </choice>
  </persName>
  <idno type="Wikidata">Q2833817</idno>
  <ptr type="url" subtype="Wikipèdia" target="https://oc.wikipedia.org/wiki/Alexandre_lo_Grand"></ptr>
  <note type="below">plan coneget en euròpa</note>
</person>

```

3.4 Les informations étymologiques

On les indique entre les balises **<etym>** et **</etym>**. Celles-ci se trouvent sous la balise **<entry>** si elles donnent des informations sur le mot vedette, sous la balise **<cit type="translation">** si elles donnent des informations sur une traduction, sous la balise **<re>** si elles donnent des informations sur le mot principal d'une sous-entrée.

La balise **<etym>** est complétée d'un attribut « **corresp** » qui permet de relier l'étymologie à la forme qu'elle complète. On y insère l'identifiant **xml:id** de la forme ou de l'étyomon concerné précédé de #. On peut avoir plusieurs identifiants à la suite.

L'étymologie peut concerter une forme (mot-vedette, traduction...) ou un étyomon d'un autre groupe d'informations étymologiques.

Exemples :

Etymologie pour une forme

```
<etym corresp="#form-15">  
...  
</etym>
```

Etymologie pour plusieurs formes

```
<etym corresp="#form-89#form-90">  
...  
</etym>
```

Etymologie pour un étyomon

```
<etym corresp="#mentioned-10">  
...  
</etym>
```

L'étymologie peut contenir les balises :

- **<bibl>** : source de l'étymologie (voir 3.2.3)
- **<gloss>** : information sur l'étyomon (traduction, transcription, définition...)
- **<lang>** : langue d'origine, indiquée par un code ISO, éventuellement dans l'attribut « **norm** »
- **<mentioned>** : étyomon
- **<note>** : informations diverses
- **<usg>** : informations d'usage de l'étyomon (voir 3.2.9)

La balise **<mentioned>** est complétée par un attribut « **xml:id** » qui prend la forme « **mentioned-[numero]** ».

La balise **<gloss>** est complétée par l'attribut « **type** », indiquant le type d'information donnée sur l'étyomon (**definition**, **transcription**, **translation**, **type**, **divers**) et, s'il y a un ou plusieurs **<mentioned>** par l'attribut « **corresp** », qui indique de quel étyomon il est la glose (on donne la valeur de l'attribut « **xml:id** » de l'étyomon précédent de #).

Pour indiquer les valeurs des balises **<gloss>** et **<mentioned>**, on utilise les balises **<reg>**, **<orig>** et **<expan>** (cf. 4).

La balise **<lang>** porte les attributs « **norm** », « **orig** » et « **expand** » (cf. 4).

Exemples :

<i>Original : « De l'occ. "espic" : épi »</i>
<i>Affiché (éd. papier) : « De l'oc. "espic" (épi) »</i>
<i>Affiché (éd. numérique) : « De l'occitan "espic" (épi) »</i>
<pre><etym corresp="#form-53"> <lang norm="oc-grclass" orig="De l'occ." expand="De l'occitan">De l'oc.</lang> <mentioned xml:id="mentioned-65"> <choice> <orig>espic</orig> <reg>espic</reg> </choice> </mentioned> <gloss type="translation" corresp="#mentioned-65"> <choice> <orig>épi</orig> <reg>épi</reg> </choice> </gloss> </etym></pre>

<i>Original : « mot gasc. selon Palay »</i>
<i>Affiché (éd. papier et numérique) : « mot gascon (Palay) »</i>
<pre><etym corresp="#form-212"> <lang norm="oc-gascon-grclass" orig="mot gasc.">mot gascon</lang> <bibl> <note rend="invisible" place="above">selon</note> <author> <choice> <orig>Palay</orig> <reg>Palay</reg> </choice> </author> </bibl> </etym></pre>

<i>Original : « du latin pyramida emprunté au grec ancien pyramis »</i>
<i>Affiché (éd. papier et numérique) : « latin pyramida ; issu du grec ancien pyramis »</i>
<pre><etym corresp="#form-30"> <lang norm="la" orig="du latin">latin</lang></pre>

```

<mentioned xml:id="mentioned-12">
  <choice>
    <orig>pyramida</orig>
    <reg>pyramida</reg>
  </choice>
</mentioned>
</etym>
<etym corresp="#mentioned-12">
  <lang norm="grc" orig="emprunté au grec ancien">issu du grec ancien</lang>
  <mentioned xml:id="mentioned-13">
    <choice>
      <orig>pyramis</orig>
      <reg>pyramis</reg>
    </choice>
  </mentioned>
</etym>

```

Affiché (éd. papier et numérique) : « onomatopée, cri du canard »

```

<etym corresp="#form-521">
  <gloss type="type">
    <choice>
      <reg>onomatopée</reg>
    </choice>
  </gloss>
  <gloss type="definition">
    <choice>
      <reg>cri du canard</reg>
    </choice>
  </gloss>
</etym>

```

3.5 Les différentes acceptions

Les différentes définitions, traductions, sont regroupées dans une ou plusieurs acceptances encadrées par les balises **<sense>** et **</sense>**.

Pour faciliter l'utilisation par des outils automatiques, les sens qui étaient imbriqués dans les dictionnaires papier ont été désimbriqués. Mais il est possible de garder une trace de l'ancienne imbrication grâce à l'attribut « **level** ». Il prend comme valeur un nombre qui correspond au niveau de l'acceptation dans l'imbrication des acceptations, 1 étant le niveau le moins imbriqué. Par exemple, l'acceptation I aura pour niveau 1, l'acceptation I-A aura pour niveau 2, l'acceptation I-A-1 aura pour niveau 3... Dans une entrée sans imbrication des acceptations, la valeur sera toujours de 1.

La balise **<sense>** accepte aussi un attribut facultatif « **n** » qui indique la numérotation de l'acceptation dans l'affichage final ou original (ex. : II, a, 1.3...).

Les acceptances possèdent un identifiant unique « **xml:id** » de type : **sense-1, sense-2, sense-3...**

Exemple :

<i>Original :</i>
I -

```

1. texte1
2. texte2
// - texte3
///-
1. texte4
2.
  a) texte5
  b) texte6

```

Affiché (éd. papier et numérique) :

- 1.1 : texte1
- 1.2 : texte2
- 2 : texte3
- 3.1 : texte4
- 3.1.1 : texte5
- 3.2.2 : texte6

```

<sense xml:id="sense-1" level="2" n="1.1">texte1</sense>
<sense xml:id="sense-2" level="2" n="1.2">texte2</sense>
<sense xml:id="sense-3" level="1" n="2">texte3</sense>
<sense xml:id="sense-4" level="2" n="3.1">texte4</sense>
<sense xml:id="sense-5" level="3" n="3.2.1">texte5</sense>
<sense xml:id="sense-6" level="3" n="3.2.2">texte6</sense>

```

Si le dictionnaire est multilingue ou multivariété, on peut avoir besoin d'indiquer qu'une acceptation d'un mot ne concerne que certaines langues/variétés dans lesquelles le mot-vénette existe. On peut l'indiquer grâce à l'attribut « **xml:lang** ». Il prend en valeur un ou plusieurs codes langues de l'IANA (voir en 7.1) séparés par « / ».

Affiché (éd. papier et numérique) :

casa gasc., leng., prov.

1. case
2. (gasc., leng.) maison

```

<entry xml:id="413">
  <form subtype="word" type="lemma" xml:lang="oc-gascon-grclass/oc-lengadoc-grclass/oc-provenc-
  grclass" n="1" xml:id="form-275" cert="ok" value="casa">
    <orth norm="casa" value="casa">casa</orth>
    <gramGrp>
      <gram type="catDet" norm="NomComFS"></gram>
      <gram type="catGen" norm="NomComF"></gram>
    </gramGrp>
  </form>
  <sense xml:id="sense-408" level="1">
    <cit type="translation">
      <form subtype="word" type="lemma" xml:lang="fr" xml:id="form-276" n="1" cert="ok" value="case">
        <orth norm="case" value="case" orig="casa">case</orth>
        ...
      </form>
    </cit>
  </sense>
  <sense xml:id="sense-409" level="1" xml:lang="oc-gascon-grclass/oc-lengadoc-grclass">
    <cit type="translation">
      <form subtype="word" type="lemma" xml:lang="fr" xml:id="form-277" n="1" cert="ok" value="maison">
        <orth norm="maison" value="maison" orig="casa">maison</orth>
      </form>
    </cit>
  </sense>

```

```
...  
</form>  
</cit>  
</sense>  
</entry>
```

3.5.1 Les informations d'usage

Sous les balises **<sense>**, on peut trouver entre les balises **<usg>** et **</usg>** les informations d'usage indiquant à quelles utilisations du mot correspond cette acceptation (voir 3.2.9).

3.5.2 La définition

On peut trouver entre les balises **<def>** et **</def>** une définition du mot-vedette.

L'attribut « **xml:id** » est l'identifiant unique de la définition. Il prend comme valeur « **def-[nombre]** »

L'attribut « **xml:lang** » indique, grâce aux codes langues de l'IANA, dans quelle langue est la définition. Si la définition concerne plusieurs langues et/ou plusieurs variétés, on les indique toutes dans l'attribut « **xml:lang** » en les séparant par « / ».

La balise **<def>** porte les attributs « **norm** », « **orig** » et « **expand** » (cf. 4).

Exemple :

Affiché (éd. papier et numérique) : « *Longue plume à l'aile ou à la queue d'un oiseau.* »

```
<def xml:id="def-1" xml:lang="fr" norm="Longue plume à l'aile ou à la queue d'un  
oiseau.">Longue plume à l'aile ou à la queue d'un oiseau.</def>
```

Affiché (éd. papier) : « *Mot inventé par F. Mistral pour désigner une abeille.* »

Affiché (éd. numérique) : « *Mot inventé par Frédéric Mistral pour désigner une abeille.* »

```
<def xml:id="def-1" xml:lang="fr" norm="Mot inventé par Frédéric Mistral pour désigner une  
abeille." expand="Mot inventé par Frédéric Mistral pour désigner une abeille.">Mot inventé par  
F. Mistral pour désigner une abeille.</def>
```

3.5.3 Les traductions et références

Les traductions du mot-vedette sont indiquées entre les balises **<cit type="translation">** et **</cit>**. Elles sont comprises directement sous l'acceptation **<sense>**. Il peut s'agir de formes **<form>** ou de personnes **<person>**.

Les références sont indiquées entre les balises **<cit type="reference">** et **</cit>**. Elles sont comprises directement sous les renvois **<xref>** (voir en 3.5.4).

A l'intérieur des balises **<cit>** et **</cit>**, les mots traduisant le mot-vedette ou renvoyant vers une forme principale sont indiqués entre les balises **<form>** et **</form>** et peuvent contenir les mêmes balises qu'un mot-vedette (voir 3.2). Une traduction **<cit>** peut contenir plusieurs formes **<form>**, si elles sont équivalentes (variantes).

Les traductions et références peuvent également comprendre les balises **<etym>** et **</etym>** (voir 3.4) pour indiquer les informations étymologiques sur l'une des formes citées.

Elles peuvent contenir des informations d'usage entre les balises **<usg>** et **</usg>** (voir 3.2.9), ou diverses informations entre les balises **<note>** et **</note>**.

Exemples :

Affiché (éd. papier et numérique) : « halèra (qui essouffle) »

```
<cit type="translation">
  <form subtype="word" type="lemma" xml:lang="oc-gascon-grclass" xml:id="form-712" n="1" cert="ok"
value="halèra">
    <orth norm="halèra" value="halèra">halèra</orth>
  </form>
  <usg type="hint" norm="qui essouffle"></usg>
  ...
</cit>
```

Original : « agusatge / agusada (f) »

Affiché (éd. papier et numérique) : « agusatge / agusada (f.) »

```
<cit type="translation">
  <form subtype="word" type="lemma" xml:lang="oc-gascon-grclass" xml:id="form-15" n="1" cert="ok"
value="agusatge">
    <orth norm="agusatge" value="agusatge" orig="agusatge">agusatge</orth>
    ...
  </form>
  <form subtype="word" type="lemma" xml:lang="oc-gascon-grclass" xml:id="form-16" n="1" cert="ok"
value="agusada">
    <orth norm="agusada" value="agusada" orig="agusada">agusada</orth>
    <gramGrp>
      <gen norm="feminine" orig="f">f.</gen>
    </gramGrp>
    ...
  </form>
</cit>
```

3.5.4 Les renvois

Les renvois (tout ce qui renvoie à une autre forme dans le document ou dans un autre document) sont indiquées entre les balise **<xr>** et **</xr>**.

La balise a un attribut « type » qui indique le type des références qu'il contient. Il prend les valeurs suivantes :

- **cf** : renvois
- **abbreviations** : formes abrégées
- **antonyms** : antonymes
- **augmentatives** : augmentatifs
- **binName** : dénomination binomiale (nom de genre + nom d'espèce) pour une plante ou un animal
- **cognates** : mot équivalent dans une autre langue (qui n'est pas forcément une traduction)
- **contractions** : formes contractées

- **demonyms** : nom des habitants (pour une ville, un pays...)
- **deriveds** : dérivés
- **diminutives** : diminutifs
- **full** : forme pleine (pour les abréviations, les symboles, les sigles...)
- **lemma** : lemme du mot (dans des tournures comme « part. pass. de "cantar" »)
- **sigle** : sigle
- **symbol** : symbole (ex. : \$ pour dollar)
- **synonyms** : synonymes
- **hyperonyms** : hyperonymes
- **keywords** : mots-clés
- **variants** : variantes

Cette liste peut être complétée.

L'affichage de ce type se fait via une sous-balise obligatoire **<lbl type="reference"></lbl>** qui prend les attributs « **norm** », « **orig** » et « **expand** » (voir en 4).

A l'intérieur des balises, chaque référence est encadrée par les balises **<cit type="reference">** et **</cit>** qui fonctionnent comme les **<cit type="translation">** (voir en 3.5.3).

Dans le cas de renvoi vers des personnes-vedettes, on utilise les balises **<person>** (voir en 3.3) directement sous la balise **<xr>**.

Dans le cas d'un renvoi vers un ouvrage complet, on utilise les balises **<bibl></bibl>** (voir en 3.2.3) sous les balises **<xr type="cf"></xr>**.

Les balises **<xr></xr>** peuvent également contenir des sous-balises **<note>** et **</note>**.

Exemples :

<i>Original : « Voir aussi : batailleur, bagarreur »</i>
--

<i>Affiché (éd. papier) : « Cf. : bagarreur »</i>

<i>Affiché (éd. numérique) : « Voir : bagarreur »</i>

<pre><xr type="cf"> <lbl type="reference" norm="cf" orig="Voir aussi" expand="Voir">Cf.</lbl> <cit type="reference"> <form subtype="word" type="lemma" xml:lang="fr" cert="current" xml:id="form-350" n="1" value="bagarreur"> <orth norm="bagarreur" value="bagarreur" orig="bagarreur">bagarreur</orth> <gramGrp> <gram type="catDet" norm="AdjMS"></gram> <gram type="catGen" norm="Adj"></gram> </gramGrp> </form> </cit> </xr></pre>

<i>Affiché (éd. papier et numérique) : « Syn. : polit, polida »</i>

<pre><xr type="synonyms"> <lbl type="reference" norm="synonyms">Syn.</lbl> <cit type="reference"></pre>

```

<form subtype="word" type="lemma" xml:lang="oc-gascon-grclass" cert="current" xml:id="form-350"
n="1" value="polit">
    <orth norm="polit" value="polit">polit</orth>
    <gramGrp>
        <gram type="catDet" norm="AdjMS"></gram>
        <gram type="catGen" norm="Adj"></gram>
    </gramGrp>
    <form subtype="word" type="inflected" xml:lang="oc-gascon-grclass" cert="current" xml:id="form-351"
n="1" value="polida">
        <orth norm="polida" value="polida">polida</orth>
        <gramGrp>
            <gram type="catDet" norm="AdjFS"></gram>
            <gram type="catGen" norm="Adj"></gram>
        </gramGrp>
    </form>
</form>
</cit>
</xr>

```

Original : « En gascon : Que vau au castèth »

Affiché (éd. papier et numérique) : « Var. : Que vau au castèth »

```

<xr type="variants">
    <note rend="invisible" place="above">En gascon :</note>
    <lbl type="reference" norm="variants">Var.</lbl>
    <cit type="reference">
        <form subtype="phrase" type="lemma" xml:lang="oc-gascon-grclass" cert="current" xml:id="form-350"
n="1" value="Que vau au castèth">
            <orth norm="Que vau au castèth" value="Que vau au castèth" orig="Que vau au castèth">Que vau au
            castèth</orth>
            <gramGrp>
                <gram type="catDet" norm="AdjMS"></gram>
                <gram type="catGen" norm="Adj"></gram>
            </gramGrp>
        </form>
    </cit>
</xr>

```

Original : « concernant la variété pyrénéenne >> isard »

Affiché (éd. papier et numérique) : « Cf. : isard (concernant la variété pyrénéenne) »

```

<xr type="cf">
    <note rend="visible" place="above">concernant la variété pyrénéenne</note>
    <lbl type="reference" norm="cf">Cf.</lbl>
    <cit type="reference">
        <form subtype="phrase" type="lemma" xml:lang="fr" cert="current" xml:id="form-350" n="1"
value="isard">
            <orth norm="isard" value="isard" orig="isard">isard</orth>
            <gramGrp>
                <gram type="catDet" norm="NomComMS"></gram>
                <gram type="catGen" norm="NomComM"></gram>
            </gramGrp>
        </form>
    </cit>
</xr>

```

Original : « voir épinoche (sens principal) »

Affiché (éd. papier et numérique) : « cf. : épinoche (sens principal) »

```
<xr type="cf">
<lbl type="reference" norm="cf" orig="voir">Cf.</lbl>
<cit type="reference">
<form subtype="phrase" type="lemma" xml:lang="fr" cert="current" xml:id="form-350" n="1"
value="épinoche">
<orth norm="épinoche" value="épinoche" orig="épinoche">épinoche</orth>
<gramGrp>
<gram type="catDet" norm="NomComFS"></gram>
<gram type="catGen" norm="NomComF"></gram>
</gramGrp>
</form>
<usg type="plev" norm="usually" orig="sens principal">sens principal</usg>
</cit>
</xr>
```

Original : « Voir Laux »

Affiché (éd. papier et numérique) : « cf. : Laux, 2004 »

```
<xr type="cf">
<lbl type="reference" norm="cf" orig="Voir">Cf.</lbl>
<bibl>
<author key="Christian Laux">
<choice>
<reg>Laux</reg>
<orig>Laux<orig>
</choice>
</author>
<date when="2004">
<choice>
<reg>2004</reg>
</choice>
</date>
</bibl>
</xr>
```

*Affiché (éd. papier et numérique) : « cf. : Padena (1944-2021),
Coluche (1944-1986) »*

```
<xr type="cf">
<lbl type="reference" norm="cf">Cf.</lbl>
<person xml:id="person-494" sortkey="padena">
<persName>
<addName sort="1" type="stagename">Padena</addName>
<choice>
<reg>Padena</reg>
</choice>
</persName>
<birth>
<date when="1944">1944</date>
</birth>
<death>
<date when="2021">2021</date>
```

```

</death>
</person>
<person xml:id="person-495" sortkey="coluche">
  <persName>
    <addName sort="1" type="stagename">Coluche</addName>
    <choice>
      <reg>Coluche</reg>
    </choice>
  </persName>
  <birth>
    <date when="1944">1944</date>
  </birth>
  <death>
    <date when="1986">1986</date>
  </death>
</person>
</xr>

```

Affiché (éd. papier) : « *vòli* : 1. pp de "voler" 2. pp de "volar" »

Affiché (éd. numérique) : « *vòli* : 1. participe passé de "voler" 2. participe passé de "volar" »

```

<entry xml:id="entry-1258">
  <form subtype="word" type="inflected" xml:lang="oc-gascon-grclass"
  xml:id="form-498" n="1" cert="current" value="vòli">
    <orth norm="vòli" value="vòli">vòli</orth>
    <gramGrp>
      <gram type="catDet" norm="VerbeIndPres1s"></gram>
      <gram type="catGen" norm="Verbe"></gram>
    </gramGrp>
  </form>
  <sense xml:id="sense-542" level="1">
    <xr type="lemma">
      <lbl type="reference" norm="VerbeIndPres1s" expand="participe passé
de">pp de</lbl>
      <cit type="reference">
        <form subtype="word" type="lemma" xml:lang="fr" xml:id="form-1259"
n="1" cert="current" value="voler">
          <orth norm="voler" value="voler">voler</orth>
          <gramGrp>
            <gram type="catDet" norm="VerbeInf"></gram>
            <gram type="catGen" norm="Verbe"></gram>
          </gramGrp>
        </form>
      </cit>
    </xr>
  </sense>
  <sense xml:id="sense-543" level="1">
    <xr type="lemma">
      <lbl type="reference" norm="VerbeIndPres1s" expand="participe passé
de">pp de</lbl>
      <cit type="reference">
        <form subtype="word" type="lemma" xml:lang="fr" xml:id="form-1260"
n="1" cert="current" value="volar">
          <orth norm="volar" value="volar">volar</orth>
          <gramGrp>
            <gram type="catDet" norm="VerbeInf"></gram>

```

```

<gram type="catGen" norm="Verbe"></gram>
</gramGrp>
</form>
</cit>
</xr>
</sense>
</entry>
```

3.5.5 Les sous-entrées

Les entrées relatives au mot-vedette (expressions, exemples, citations, mots composés, dérivés...) ne se trouvant pas dans une entrée indépendante sont placées entre les balises **<re>** et **</re>**, sous les balises **<sense>** ou directement sous la balise **<entry>**. Elles sont complétées par l'attribut « **type** », qui prend les valeurs suivantes :

- **attestation** : attestations historiques du mot dans des citations d'œuvres anciennes
- **citation** : citations contenant le mot-vedette
- **collocation** : mot-vedette avec une ou plusieurs collocations (ex. : « inciter à » pour « inciter »).
- **compound** : mot composé (en un seul mot ou avec un tiret)
- **demonym** : démonyme (nom d'habitant d'un lieu qui est le mot-vedette)
- **derived** : dérivé
- **example** : exemple
- **expression** : expression
- **interjection** : interjection
- **locution** : locution
- **proper** : nom propre
- **pronominal** : dérivé pronominal d'un verbe non pronominal (ex. : « s'abonar » pour « abonar »)
- **proverb** : proverbe, dicton
- **unidentified** : type non identifié

Elles ont un attribut « **xml:id** » qui prend en valeur un identifiant unique de type « **re-[nombre]** ».

Elles sont composées comme les entrées principales, et peuvent contenir les balises suivantes :

- **<form>** : voir 3.2
- **<person>** : voir 3.3
- **<etym>** : voir 3.4
- **<sense>** : 3.5 (ne pouvant pas contenir eux-même une sous-balise **<re>**)
- **<note>** : informations diverses

Elles peuvent aussi comporter une sous-balise de type **<lbl type="subentry"></lbl>** qui indique le type de sous-entrée, et qui prend les attributs « **norm** », « **orig** » et « **expand** » (voir en 4).

Exemples :

Original : « ~ nain (*Ulex nanus*) : toja »

Affiché (éd. papier et numérique) : « ajonc nain (*Ulex nanus*) : toja »

```

<re xml:id="re-513" type="compound">
  <form subtype="phrase" type="lemma" xml:lang="fr" xml:id="form-54" n="1" cert="current" value="ajonc
nain">
    <orth norm="ajonc nain" value="ajonc nain" orig="~ nain">ajonc nain</orth>
    ...
  </form>
  <sense xml:id="sense-65" level="1">
    <usg type="binName" norm="Ulex nanus" orig="Ulex nanus">Ulex nanus</usg>
    <cit type="translation">
      <form subtype="word" type="lemma" xml:lang="oc-gascon-grclass" xml:id="form-55" n="1" cert="ok"
value="toja">
        <orth norm="toja" value="toja" orig="toja">toja</orth>
        ...
      </form>
    </cit>
  </sense>
</re>

```

Affiché (éd. papier) : « *hocher la tête – (vert.) capejar – (horiz.) cabocejar* »

Affiché (éd. numérique) : « *hocher la tête – (verticalement) capejar – (horizontalement)
cabocejar* »

```

<re xml:id="re-129" type="unidentified">
  <form subtype="phrase" type="lemma" xml:lang="fr" xml:id="form-11" n="1" cert="current" value="hocher
la tête">
    <orth norm="hocher la tête" value="hocher la tête">hocher la tête</orth>
    ...
  </form>
  <sense xml:id="sense-13" level="1">
    <usg type="hint" norm="verticalement" expand="verticalement">vert.</usg>
    <cit type="translation">
      <form subtype="word" type="lemma" xml:lang="oc-gascon-grclass" xml:id="form-12" n="1" cert="ok"
value="capejar">
        <orth norm="capejar" value="capejar">capejar</orth>
        ...
      </form>
    </cit>
  </sense>
  <sense xml:id="sense-14" level="1">
    <usg type="hint" norm="horizontalement" expand="horizontalement">horiz.</usg>
    <cit type="translation">
      <form subtype="word" type="lemma" xml:lang="oc-gascon-grclass" xml:id="form-13" n="1" cert="ok"
value="cabocejar">
        <orth norm="cabocejar" value="cabocejar">cabocejar</orth>
        ...
      </form>
    </cit>
  </sense>
</re>

```

Affiché (éd. papier et numérique) : « *gentilé : Paulin* »

```

<re xml:id="re-3398" type="demonym">
  <lbl type="subentry" norm="gentilé">gentilé</lbl>
  <sense level="1" xml:id="sense-33400">
    <cit type="translation">

```

```

<form subtype="word" xml:lang="fr" xml:id="form-15831" n="1" cert="current" value="Paulin">
  <orth norm="Paulin" value="Paulin">Paulin</orth>
  ...
</form>
</cit>
</sense>
</re>

```

3.6 Indiquer un fichier lié

On peut relier un ou plusieurs fichiers à une entrée grâce à la balise **<ptr>**. Elle est complétée par un attribut « **type** » qui indique le type de contenu du fichier. Il peut prendre les valeurs suivantes :

- **audio** : fichiers audios
- **img** : images
- **uri** : identifiant uniforme de ressource (par exemple identifiant Wikidata)
- **url** : lien vers un site web
- **video** : fichier vidéo
- **other** : tous les autres fichiers

L'attribut « **target** » donne le nom ou l'url complète du fichier lié.

On peut aussi ajouter un attribut « **subtype** » si on a plusieurs fichiers de même type destinés à des utilisations différentes. Cet attribut accepte n'importe quelle valeur.

On peut enfin indiquer un attribut « **targetLang** » qui indique la langue du contenu ciblé par le lien. Il prend comme valeur un code langue de l'IANA (voir en 7.1).

Les fichiers liés peuvent être indiqués sous l'entrée ou sous une forme.

Exemples :

Affiché (éd. papier et numérique) : « Aver la pluma lusenta - Avoir bonne mine »
<pre> <entry xml:id="entry-458"> <ptr type="img" target="plumas.jpg"></ptr> <ptr type="url" target="https://www.locongres.org/expressoc/10211-aver-la-pluma-lusenta" targetLang="oc-grclass"></ptr> <ptr type="url" target="https://www.locongres.org/expressoc/10211-avoir-bonne-mine" targetLang="fr"></ptr> <form subtype="phrase" type="lemma" xml:lang="oc-gascon-grclass" xml:id="form-498" n="1" cert="current" value="Aver la pluma lusenta"> <orth norm="Aver la pluma lusenta" value="Aver la pluma lusenta">Aver la pluma lusenta</orth> <ptr type="sound" target="plumalusenta.mp3"></ptr> ... </form> <sense xml:id="sense-450" level="1"> <cit type="translation"> <form subtype="phrase" type="lemma" xml:lang="fr" xml:id="form-499" n="1" cert="current" value="Avoir bonne mine"> <orth norm="Avoir bonne mine" value="Avoir bonne mine">Avoir bonne mine</orth> ... </form> </cit> </sense> </entry> </pre>

Affiché (éd. papier et numérique) : « Haget - Hagetmau »

```
<entry xml:id="entry-53">
  <ptr type="img" target="hagetmau.jpg"></ptr>
  <ptr type="url" subtype="congres" target="https://www.locongres.org/topoc/9903-haget"></ptr>
  <ptr type="url" subtype="wikipedia" target="https://oc.wikipedia.org/wiki/Hagetmau"
    targetLang="fr"></ptr>
  <ptr type="url" subtype="wikipedia" target="https://oc.wikipedia.org/wiki/Haget" targetLang="oc-
    grclass"></ptr>
    <form subtype="word" type="lemma" xml:lang="oc-gascon-grclass" xml:id="form-500" n="1"
      cert="current" value="Haget">
      <orth norm="Haget" value="Haget">Haget</orth>
      ...
    </form>
    <sense xml:id="sense-360" level="1">
      <cit type="translation">
        <form subtype="word" type="lemma" xml:lang="fr" xml:id="form-501" n="1" cert="current"
          value="Hagetmau">
          <orth norm="Hagetmau" value="Hagetmau">Hagetmau</orth>
          ...
        </form>
      </cit>
    </sense>
  </entry>
```

Lorsqu'il s'agit d'URI, on peut également trouver la balise **<ptr>** directement sous une acceptation **<sense>**. Cela permet, notamment, de lier une acceptation d'un mot au concept qu'il représente.

Exemple :

chien : 1 - (*animal*) 2 - (*d'une arme à feu*)

```
chien : 1 - (animal) 2 - (d'une arme à feu)
<entry xml:id="entry-2569">
  <form subtype="word" type="lemma" xml:lang="fr" xml:id="form-45841" n="1" cert="current"
    value="chien">
    <orth norm="chien" value="chien">chien</orth>
    ...
  </form>
  <sense xml:id="sense-32569" level="1">
    <ptr type="uri" subtype="Wikidata" target="http://www.wikidata.org/entity/Q144"></ptr>
    <usg type="hyper" norm="animal">animal</usg>
    ...
  </sense>
  <sense xml:id="sense-32570" level="1">
    <ptr type="uri" subtype="Wikidata" target="http://www.wikidata.org/entity/Q2333657"></ptr>
    <usg type="hint" norm="d'une arme à feu">d'une arme à feu</usg>
    ...
  </sense>
</entry>
```

3.7 La balise **<note>**

Elle peut se trouver sous toutes les balises ne pouvant contenir ni les balises **<reg>**, **<orig>** et **<expan>**, ni du texte brut, et encadre toutes les informations qui ne correspondent à aucune des autres balises (cf 3.2.11).

4 Les balises `<reg>`, `<orig>` et `<expan>` et les attributs « norm », « orig » et « expand »

Dans les dictionnaires finaux, les informations ont quatre facettes :

- La version telle qu'elle était donnée dans le dictionnaire d'origine
- La version normalisée pour les ordinateurs
- La version telle qu'on veut l'afficher dans le dictionnaire de sortie à destination d'humains, dans une édition papier
- La version telle qu'on veut l'afficher dans le dictionnaire de sortie à destination d'humains, dans une édition numérique (où il y a davantage de place)

Par exemple, le genre d'un nom pouvait être indiqué dans le dictionnaire d'origine grâce à l'abréviation « fem. ». Dans le cadre de l'harmonisation de l'affichage des dictionnaires du Congrès, nous voulons que le genre soit indiqué avec l'abréviation « f. » dans une édition papier où la place est comptée, et « fem. » dans une édition numérique où le nombre de caractères importe peu. Et la forme normalisée du genre, utilisée pour le traitement automatique, est « feminine ».

On peut indiquer ceci de deux façons, en fonction de la balise concernée :

- Grâce aux attributs : « **orig** » indique quelle était la forme dans le dictionnaire d'origine, « **norm** » indique la forme normalisée, le contenu entre les balises indique la forme telle qu'elle doit être affichée dans une édition papier et l'attribut « **expand** » indique la forme à afficher dans une édition numérique. L'attribut « **norm** » est obligatoire. S'il n'y a pas de contenu entre les balises, rien n'est affiché. S'il n'y a pas d'attribut « **orig** », on considère que le dictionnaire d'origine ne comportait pas l'information. S'il n'y a pas d'attribut « **expand** », on affiche pour l'édition numérique la même chose que pour l'édition papier (c'est à dire le contenu entre les balises).
- Grâce aux sous-balises : à l'intérieur d'une balise `<choice></choice>`, on met la forme du dictionnaire d'origine entre les balises `<orig></orig>`, la forme à afficher pour une édition papier entre les balises `<reg></reg>` et la forme à afficher pour une édition numérique entre les balises `<expan></expan>`. S'il y a une normalisation du champ, la forme normalisée est indiquée, en général, dans un attribut (« **key** », « **when** »...). S'il n'y a pas de balise `<orig>`, on considère que le dictionnaire d'origine ne comportait pas l'information. S'il n'y a pas de balise `<expan>`, dans une édition numérique, on affichera la même chose que dans l'édition papier (le contenu de la balise `<reg>`).

Exemples :

Original : « fém »
Affiché (éd. papier) : « f. »
Affiché (éd. numérique) : « fémin. »
<code><lbl xml:lang="fr" type="catDet" norm="feminine" orig="fém" expand="fémin.">f.</lbl></code>

Original : « Laus »
Affiché (éd. papier) : « C. Laux »

Affiché (éd. numérique) : « Christian Laux »

```
<author key="Christian Laux">
<choice>
<orig>Laus</orig>
<reg>C. Laux</reg>
<expan>Christian Laux</expan>
</choice>
</author>
```

Pas d'original

Affiché (éd. papier) : « fig. »

Affiché (éd. numérique) : « fig. »

```
<usg type="style" norm="figurative">fig.</usg>
```

Pas d'original

Affiché (éd. papier) : « 12/06/1994 »

Affiché (éd. numérique) : « 12/06/1994 »

```
<date when="1994">
<choice>
<reg>12/06/1994</reg>
</choice>
</date>
```

5 Normes pour les catégories grammaticales

Les catégories grammaticales sont en général indiquées avec deux étiquettes :

- la catégorie générale, correspondant au PoS que partagent en général un lemme et ses formes fléchies.
- la catégorie détaillée, correspondant au PoS + les informations morphosyntaxiques (genre, nombre, informations de conjugaison...).

Dans certains lexiques partagés avec des partenaires, on peut adjoindre à ces étiquettes des informations grammaticales qui suivent la norme internationale *Eagles* ou *Grace*.

Dans d'autres lexiques, notamment ceux à destination du machine learning, on peut ajouter une étiquette plus générale de famille grammaticale. Elle rassemble les catégories générales ayant un fonctionnement très similaire.

5.1 Jeu d'étiquettes du Congrès

Pour étiqueter ses données, le Congrès utilise des étiquettes

5.1.1 Verbes

Les verbes ont pour catégorie générale **Verbe**. Leurs catégories détaillées comportent les informations de temps, de mode et de personne.

Ils font partie de la famille grammaticale **Verbe**.

Catégories détaillées étiquetant des verbes à l'indicatif présent

Catégorie détaillée	Formes fléchies concernées
VerbelndPres1s	1re personne du singulier au présent de l'indicatif
VerbelndPres2s	2e personne du singulier au présent de l'indicatif
VerbelndPres3s	3e personne du singulier au présent de l'indicatif
VerbelndPres1p	1re personne du pluriel au présent de l'indicatif
VerbelndPres2p	2e personne du pluriel au présent de l'indicatif
VerbelndPres3p	3e personne du pluriel au présent de l'indicatif

Catégories détaillées étiquetant des verbes à l'indicatif imparfait

Catégorie détaillée	Formes fléchies concernées
Verbelndlmp1s	1re personne du singulier de l'imparfait de l'indicatif
Verbelndlmp2s	2e personne du singulier de l'imparfait de l'indicatif
Verbelndlmp3s	3e personne du singulier de l'imparfait de l'indicatif
Verbelndlmp1p	1re personne du pluriel de l'imparfait de l'indicatif
Verbelndlmp2p	2e personne du pluriel de l'imparfait de l'indicatif
Verbelndlmp3p	3e personne du pluriel de l'imparfait de l'indicatif

Catégories détaillées étiquetant des verbes à l'indicatif présent

Catégorie détaillée	Formes fléchies concernées
VerbelndPret1s	1re personne du singulier du présent de l'indicatif
VerbelndPret2s	2e personne du singulier du présent de l'indicatif
VerbelndPret3s	3e personne du singulier du présent de l'indicatif
VerbelndPret1p	1re personne du pluriel du présent de l'indicatif
VerbelndPret2p	2e personne du pluriel du présent de l'indicatif
VerbelndPret3p	3e personne du pluriel du présent de l'indicatif

Catégories détaillées étiquetant des verbes à l'indicatif futur

Catégorie détaillée	Formes fléchies concernées
VerbelndFut1s	1re personne du singulier du futur de l'indicatif
VerbelndFut2s	2e personne du singulier du futur de l'indicatif
VerbelndFut3s	3e personne du singulier du futur de l'indicatif
VerbelndFut1p	1re personne du pluriel du futur de l'indicatif
VerbelndFut2p	2e personne du pluriel du futur de l'indicatif
VerbelndFut3p	3e personne du pluriel du futur de l'indicatif

Catégories détaillées étiquetant des verbes au subjonctif présent

Catégorie détaillée	Formes fléchies concernées
VerbeSubjPres1s	1re personne du singulier du présent du subjonctif
VerbeSubjPres2s	2e personne du singulier du présent du subjonctif
VerbeSubjPres3s	3e personne du singulier du présent du subjonctif
VerbeSubjPres1p	1re personne du pluriel du présent du subjonctif
VerbeSubjPres2p	2e personne du pluriel du présent du subjonctif
VerbeSubjPres3p	3e personne du pluriel du présent du subjonctif

Catégories détaillées étiquetant des verbes au subjonctif imparfait

Catégorie détaillée	Formes fléchies concernées
VerbeSubjImp1s	1re personne du singulier de l'imparfait du subjonctif
VerbeSubjImp2s	2e personne du singulier de l'imparfait du subjonctif
VerbeSubjImp3s	3e personne du singulier de l'imparfait du subjonctif
VerbeSubjImp1p	1re personne du pluriel de l'imparfait du subjonctif
VerbeSubjImp2p	2e personne du pluriel de l'imparfait du subjonctif
VerbeSubjImp3p	3e personne du pluriel de l'imparfait du subjonctif

Catégories détaillées étiquetant des verbes au conditionnel présent

Catégorie détaillée	Formes fléchies concernées
VerbeCondPres1s	1re personne du singulier du présent du conditionnel
VerbeCondPres2s	2e personne du singulier du présent du conditionnel
VerbeCondPres3s	3e personne du singulier du présent du conditionnel
VerbeCondPres1p	1re personne du pluriel du présent du conditionnel
VerbeCondPres2p	2e personne du pluriel du présent du conditionnel
VerbeCondPres3p	3e personne du pluriel du présent du conditionnel

Catégories détaillées étiquetant des verbes à l'impératif affirmatif

Catégorie détaillée	Formes fléchies concernées
VerbelImpPresAff2s	2e personne du singulier de l'impératif présent affirmatif
VerbelImpPresAff1p	1re personne du pluriel de l'impératif présent affirmatif
VerbelImpPresAff2p	2e personne du pluriel de l'impératif présent affirmatif

Catégories détaillées étiquetant des verbes à l'impératif négatif

Catégorie détaillée	Formes fléchies concernées
VerbelImpPresNeg2s	2e personne du singulier de l'impératif présent négatif
VerbelImpPresNeg1p	1re personne du pluriel de l'impératif présent négatif
VerbelImpPresNeg2p	2e personne du pluriel de l'impératif présent négatif

Catégories détaillées étiquetant des verbes au participe passé

Catégorie détaillée	Formes fléchies concernées
VerbePpMS	participe passé masculin singulier
VerbePpFS	participe passé féminin singulier
VerbePpMP	participe passé masculin pluriel
VerbePpFP	participe passé féminin pluriel

Catégorie détaillée étiquetant des verbes au participe présent

Catégorie détaillée	Formes fléchies concernées
VerbePPres	participe présent

Catégorie détaillée étiquetant des verbes à l'infinitif

Catégorie détaillée	Formes fléchies concernées
Verbelnf	infinitif

5.1.2 Noms

5.1.2.1 Noms communs

Les noms communs ont pour catégorie générale :

- NomComM : noms communs masculins
- NomComF : noms communs féminins
- NomComMP : noms communs masculins qui n'existent qu'au pluriel
- NomComFP : noms communs féminins qui n'existent qu'au pluriel

Ils font partie de la famille grammaticale NomCom.

Differentes catégories détaillées étiquetant des noms communs et catégories générales qui leur correspondent

Catégorie détaillée	Formes fléchies concernées	Catégorie générale
NomComMS	singulier	NomComM
NomComFS	singulier	NomComF
NomComMP	pluriel	NomComM ou NomComMP
NomComFP	pluriel	NomComF ou NomComFP

5.1.2.2 Noms propres

Les noms propres ont pour catégorie générale :

- NomPropMS : noms propres masculins singuliers
- NomPropFS : noms propres féminins singuliers
- NomPropMP : noms propres masculins pluriels
- NomPropFP : noms propres féminins pluriels
- NomProp : noms propres dont on ne connaît pas le genre (essentiellement les noms de ville)

Ils font partie de la famille grammaticale NomProp.

Differentes catégories détaillées étiquetant des noms propres et catégories générales qui leur correspondent

Catégorie détaillée	Formes fléchies concernées	Catégorie générale
NomProp		NomProp
NomPropFP		NomPropFP
NomPropFS		NomPropFS
NomPropMP		NomPropMP
NomPropMS		NomPropMS

5.1.3 Adjectifs qualificatifs

Ils ont pour catégorie générale Adj.

Ils font partie de la famille grammaticale Adj.

Differentes catégories détaillées étiquetant des adjectifs qualificatifs et catégorie générale qui leur correspond

Catégorie détaillée	Formes fléchies concernées	Catégorie générale
AdjMS	masculin singulier	Adj
AdjFS	féminin singulier	Adj
AdjMP	masculin pluriel	Adj
AdjFP	féminin pluriel	Adj

5.1.4 Pronoms, déterminants et adjectifs autres que qualificatifs

Tous les pronoms font partie de la famille grammaticale **Pron**. Tous les déterminants font partie de la famille grammaticale **Det**. Tous les adjectifs font partie de la famille grammaticale **Adj**.

5.1.4.1 Définis

Nous avons étiqueté en Déterminants définis (DetDef) les articles définis (« *lo tipe* », « *las peras* »).

Catégories détaillées étiquetant des déterminants définis et catégorie générale qui leur correspondent

Catégorie détaillée	Formes fléchies concernées	Catégorie générale
DetDefMS	masculin singulier	DetDef
DetDefFS	féminin singulier	DetDef
DetDefMP	masculin pluriel	DetDef
DetDefFP	féminin pluriel	DetDef

5.1.4.2 Indéfinis

Nous avons distingué :

- Les pronoms indéfinis, qui « remplacent » un nom, avec parmi eux :
 - **PronIndefS** : Les pronoms indéfinis singuliers (ex. : « *degun es pas vengut* », « *cadun n'a un* »...)
 - **PronIndefP** : Les pronoms indéfinis pluriels (ex. : « *los autres son aquí* », « *qualques uns son blaus* », « *d'unés i an pensat* »...)
 - **PronIndefNeut** : Les pronoms indéfinis neutres, qui ne portent pas d'information de genre intrinsèque (ex. : « *balha-me quicòm mai* », « *totòm o sap* », « *que demandaràs a qui que sia* »...)

Les pronoms indéfinis singuliers et pluriels, la plupart du temps, ont une forme au féminin et une forme au masculin. Les pronoms indéfinis neutres, eux, ne permettent pas de connaître le genre de ce qu'ils désignent.

- Les déterminants indéfinis (dont les articles indéfinis), qui déterminent un nom, avec parmi eux :

- **DetIndef** : les déterminants indéfinis qui se déclinent au singulier et au pluriel.
(ex. : « *una femna* », « *de pomas* », « *n'importe quel homme* », « *beaucoup de temps* »...)
- **DetIndefS** : les déterminants indéfinis qui n'existent qu'au singulier, c'est à dire toujours suivis d'un nom au singulier (ex. : « *cada jorn* », « *nat òme* », « *mai d'un còp* »...)
- **DetIndefP** : les déterminants indéfinis qui n'existent qu'au pluriel, c'est à dire toujours suivis d'un pluriel (ex. : « *plusors personas* », « *una tropèl de mainatges* »...)
- **AdjIndef** : Les adjectifs indéfinis, qui ont une fonction d'adjectif qualificatif et se trouvent en général en plus d'un déterminant (ex. : « *un autre còp* », « *lo meteis jorn* », « *un cèrt sènher Bòrdanava* »...)

Catégories détaillées étiquetant des adjectifs indéfinis et catégorie générale qui leur correspond

Catégorie détaillée	Formes fléchies concernées	Catégorie générale
AdjIndefMS	masculin singulier	AdjIndef
AdjIndefFS	féminin singulier	AdjIndef
AdjIndefMP	masculin pluriel	AdjIndef
AdjIndefFP	féminin pluriel	AdjIndef

Différentes catégories détaillées étiquetant des déterminants indéfinis et catégorie générale qui leur correspondent

Catégorie détaillée	Formes fléchies concernées	Catégorie générale
DetIndefMS	masculin singulier	DetIndef
DetIndefFS	féminin singulier	DetIndef
DetIndefMP	masculin pluriel	DetIndef
DetIndefFP	féminin pluriel	DetIndef
DetIndefSM	masculin	DetIndefS
DetIndefSF	féminin	DetIndefS
DetIndefPM	masculin	DetIndefP
DetIndefPF	féminin	DetIndefP

Différentes catégories détaillées étiquetant des pronoms indéfinis et catégories générales qui leur correspondent

Catégorie détaillée	Formes fléchies concernées	Catégorie générale
PronIndefSM	masculin	PronIndefS
PronIndefSF	féminin	PronIndefS
PronIndefPM	masculin	PronIndefP
PronIndefPF	féminin	PronIndefP

Catégorie détaillée	Formes fléchies concernées	Catégorie générale
PronIndefNeut		PronIndefNeut

5.1.4.3 Démonstratifs

Nous avons distingué :

- Les pronoms démonstratifs, qui « remplacent » un nom, avec parmi eux :
 - **PronDemNeut** : Les pronoms indéfinis neutres, qui ne portent pas d'information de genre intrinsèque (ex. : « qu'es aquò », « çò de bon »...)
 - **PronDem** : Les autres pronoms indéfinis, qui varient en fonction du genre et du nombre (ex. : « aquera n'ai pas l'auta », « aquesta, la coneissi pas »..., en français « celui-ci », « celle-là »...)
- **DetDem** : Les déterminants démonstratifs, qui déterminent un nom (ex. : « aquel arbre », « aquesta hemna »... en français « ce », « cette »...)
- **AdjDem** : Les adjectifs démonstratifs, qui qualifient un nom (ex. : « lo libre aqueste »)

Différentes catégories détaillées étiquetant des déterminants démonstratifs et catégorie générale qui leur correspondent

Catégorie détaillée	Formes fléchies concernées	Catégorie générale
DetDemMS	masculin singulier	DetDem
DetDemFS	féminin singulier	DetDem
DetDemMP	masculin pluriel	DetDem
DetDemFP	féminin pluriel	DetDem

Différentes catégories détaillées étiquetant des pronoms démonstratifs et catégories générales qui leur correspondent

Catégorie détaillée	Formes fléchies concernées	Catégorie générale
PronDemMS	masculin singulier	PronDem
PronDemFS	féminin singulier	PronDem
PronDemMP	masculin pluriel	PronDem
PronDemFP	féminin pluriel	PronDem
PronDemNeut		PronDemNeut

Différentes catégories détaillées étiquetant des adjectifs démonstratifs et catégories générales qui leur correspondent

Catégorie détaillée	Formes fléchies concernées	Catégorie générale
AdjDemMS	masculin singulier	AdjDem
AdjDemFS	féminin singulier	AdjDem
AdjDemMP	masculin pluriel	AdjDem
AdjDemFP	féminin pluriel	AdjDem

5.1.4.4 Exclamatifs

Les déterminants exclamatifs (ex. : « *quin hred !* ») ont pour catégorie générale **DetExcl**. Si besoin, les catégories générales **PronExcl** et **PronExclNeut** ont également été créées pour les pronoms exclamatifs, bien que nous n'en ayons pas rencontré jusqu'à présent.

Différentes catégories détaillées étiquetant des déterminants exclamatifs et catégorie générale qui leur correspondent

Catégorie détaillée	Formes fléchies concernées	Catégorie générale
DetExclMS	masculin singulier	DetExcl
DetExclFS	féminin singulier	DetExcl
DetExclMP	masculin pluriel	DetExcl
DetExclFP	féminin pluriel	DetExcl

Différentes catégories détaillées étiquetant des pronoms exclamatifs et catégories générales qui leur correspondent

Catégorie détaillée	Formes fléchies concernées	Catégorie générale
PronExclMS	masculin singulier	PronExcl
PronExclFS	féminin singulier	PronExcl
PronExclMP	masculin pluriel	PronExcl
PronExclFP	féminin pluriel	PronExcl
PronExclNeut		PronExclNeut

5.1.4.5 Interrogatifs

Nous avons distingué :

- Les pronoms interrogatifs, qui « remplacent » un nom, avec parmi eux :
 - **PronInterrNeut** : Les pronoms interrogatifs neutres, qui ne portent pas d'information de genre intrinsèque (ex. : « *qui t'a dit aquò ?* », « *qué ne pensas ?* »...)
 - **PronInterr** : Les autres pronoms interrogatifs, qui varient en fonction du genre et du nombre (ex. : « *lo qual es lo qu'aimas mai ?* », « *de la quala es tirat ?* »...)
- **DetInterr** : Les déterminants interrogatifs, qui déterminent un nom (ex. : « *quau poma e vòls ?* », « *qual professor as ara ?* »...)

Différentes catégories détaillées étiquetant des déterminants interrogatifs et catégorie générale qui leur correspondent

Catégorie détaillée	Formes fléchies concernées	Catégorie générale
DetInterrMS	masculin singulier	DetInterr
DetInterrFS	féminin singulier	DetInterr
DetInterrMP	masculin pluriel	DetInterr
DetInterrFP	féminin pluriel	DetInterr

Differentes catégories détaillées étiquetant des pronoms interrogatifs et catégories générales qui leur correspondent

Catégorie détaillée	Formes fléchies concernées	Catégorie générale
PronInterrMS	mASCULIN singulier	PronInterr
PronInterrMP	mASCULIN pluriel	PronInterr
PronInterrFS	fÉMININ singulier	
PronInterrFP	fÉMININ pluriel	PronInterr
PronInterrNeut		PronInterrNeut

5.1.4.6 Possessifs

Nous avons distingué :

- Les pronoms possessifs, qui « remplacent » un nom (ex. : « pren *lo mieu* », « a manjat *la soa* », « *lo tieu* es *lo melhor* »...), avec parmi eux :
 - PronPoss1s : pronoms possessifs de la première personne du singulier
 - PronPoss2s : pronoms possessifs de la deuxième personne du singulier
 - PronPoss3s : pronoms possessifs de la troisième personne du singulier
 - PronPoss1p : pronoms possessifs de la première personne du pluriel
 - PronPoss2p : pronoms possessifs de la deuxième personne du pluriel
 - PronPoss3p : pronoms possessifs de la troisième personne du pluriel
- Les déterminants possessifs, qui déterminent un nom (ex. : « *lo son* *gat* », « *ma maire* »...), avec parmi eux :
 - DetPoss1s : déterminants possessifs de la première personne du singulier
 - DetPoss2s : déterminants possessifs de la deuxième personne du singulier
 - DetPoss3s : déterminants possessifs de la troisième personne du singulier
 - DetPoss1p : déterminants possessifs de la première personne du pluriel
 - DetPoss2p : déterminants possessifs de la deuxième personne du pluriel
 - DetPoss3p : déterminants possessifs de la troisième personne du pluriel
- Les adjectifs possessifs, qui ont une fonction d'adjectif qualificatif et se trouvent en général après le verbe être (ex. : « *aquel libre* es *mieu* », « *la sòrre toa* »), avec parmi eux :
 - AdjPoss1s : adjectifs possessifs de la première personne du singulier
 - AdjPoss2s : adjectifs possessifs de la deuxième personne du singulier
 - AdjPoss3s : adjectifs possessifs de la troisième personne du singulier
 - AdjPoss1p : adjectifs possessifs de la première personne du pluriel
 - AdjPoss2p : adjectifs possessifs de la deuxième personne du pluriel
 - AdjPoss3p : adjectifs possessifs de la troisième personne du pluriel

Differentes catégories détaillées étiquetant des adjectifs possessifs et catégories générales qui leur correspondent

Catégorie détaillée	Formes fléchies concernées	Catégorie générale
AdjPoss1sMS	mASCULIN singulier	AdjPoss1s
AdjPoss1sFS	fÉMININ singulier	AdjPoss1s

AdjPoss1sMP	masculin pluriel	AdjPoss1s
AdjPoss1sFP	féminin pluriel	AdjPoss1s
AdjPoss2sMS	masculin singulier	AdjPoss2s
AdjPoss2sFS	féminin singulier	AdjPoss2s
AdjPoss2sMP	masculin pluriel	AdjPoss2s
AdjPoss2sFP	féminin pluriel	AdjPoss2s
AdjPoss3sMS	masculin singulier	AdjPoss3s
AdjPoss3sFS	féminin singulier	AdjPoss3s
AdjPoss3sMP	masculin pluriel	AdjPoss3s
AdjPoss3sFP	féminin pluriel	AdjPoss3s
AdjPoss1pMS	masculin singulier	AdjPoss1p
AdjPoss1pFS	féminin singulier	AdjPoss1p
AdjPoss1pMP	masculin pluriel	AdjPoss1p
AdjPoss1pFP	féminin pluriel	AdjPoss1p
AdjPoss2pMS	masculin singulier	AdjPoss2p
AdjPoss2pFS	féminin singulier	AdjPoss2p
AdjPoss2pMP	masculin pluriel	AdjPoss2p
AdjPoss2pFP	féminin pluriel	AdjPoss2p
AdjPoss3pMS	masculin singulier	AdjPoss3p
AdjPoss3pFS	féminin singulier	AdjPoss3p
AdjPoss3pMP	masculin pluriel	AdjPoss3p
AdjPoss3pFP	féminin pluriel	AdjPoss3p

Différentes catégories détaillées étiquetant des déterminants possessifs et catégories générales qui leur correspondent

Catégorie détaillée	Formes fléchies concernées	Catégorie générale
DetPoss1sMS	masculin singulier	DetPoss1s
DetPoss1sFS	féminin singulier	DetPoss1s
DetPoss1sMP	masculin pluriel	DetPoss1s
DetPoss1sFP	féminin pluriel	DetPoss1s
DetPoss2sMS	masculin singulier	DetPoss2s
DetPoss2sFS	féminin singulier	DetPoss2s
DetPoss2sMP	masculin pluriel	DetPoss2s
DetPoss2sFP	féminin pluriel	DetPoss2s
DetPoss3sMS	masculin singulier	DetPoss3s

DetPoss3sFS	féminin singulier	DetPoss3s
DetPoss3sMP	masculin pluriel	DetPoss3s
DetPoss3sFP	féminin pluriel	DetPoss3s
DetPoss1pMS	masculin singulier	DetPoss1p
DetPoss1pFS	féminin singulier	DetPoss1p
DetPoss1pMP	masculin pluriel	DetPoss1p
DetPoss1pFP	féminin pluriel	DetPoss1p
DetPoss2pMS	masculin singulier	DetPoss2p
DetPoss2pFS	féminin singulier	DetPoss2p
DetPoss2pMP	masculin pluriel	DetPoss2p
DetPoss2pFP	féminin pluriel	DetPoss2p
DetPoss3pMS	masculin singulier	DetPoss3p
DetPoss3pFS	féminin singulier	DetPoss3p
DetPoss3pMP	masculin pluriel	DetPoss3p
DetPoss3pFP	féminin pluriel	DetPoss3p

Differentes catégories détaillées étiquetant des pronoms possessifs et catégories générales qui leur correspondent

Catégorie détaillée	Formes fléchies concernées	Catégorie générale
PronPoss1sMS	masculin singulier	PronPoss1s
PronPoss1sFS	féminin singulier	PronPoss1s
PronPoss1sMP	masculin pluriel	PronPoss1s
PronPoss1sFP	féminin pluriel	PronPoss1s
PronPoss2sMS	masculin singulier	PronPoss2s
PronPoss2sFS	féminin singulier	PronPoss2s
PronPoss2sMP	masculin pluriel	PronPoss2s
PronPoss2sFP	féminin pluriel	PronPoss2s
PronPoss3sMS	masculin singulier	PronPoss3s
PronPoss3sFS	féminin singulier	PronPoss3s
PronPoss3sMP	masculin pluriel	PronPoss3s
PronPoss3sFP	féminin pluriel	PronPoss3s
PronPoss1pMS	masculin singulier	PronPoss1p
PronPoss1pFS	féminin singulier	PronPoss1p
PronPoss1pMP	masculin pluriel	PronPoss1p
PronPoss1pFP	féminin pluriel	PronPoss1p

PronPoss2pMS	masculin singulier	PronPoss2p
PronPoss2pFS	féminin singulier	PronPoss2p
PronPoss2pMP	masculin pluriel	PronPoss2p
PronPoss2pFP	féminin pluriel	PronPoss2p
PronPoss3pMS	masculin singulier	PronPoss3p
PronPoss3pFS	féminin singulier	PronPoss3p
PronPoss3pMP	masculin pluriel	PronPoss3p
PronPoss3pFP	féminin pluriel	PronPoss3p

5.1.4.7 Relatifs

Nous avons distingué :

- Les pronoms relatifs, qui « remplacent » un nom, avec parmi eux :
 - **PronRelNeut** : Les pronoms relatifs neutres, qui ne portent pas d'information de genre intrinsèque (ex. : « diga-me *a de qué* pensas », « que hè çò *qui* cau »...)
 - **PronRel** : Les autres pronoms relatifs, qui varient en fonction du genre et du nombre (ex. : « l'òme *del qual* parli », « la causa *a la quau e* pensas »...)
- **DetRel** : Les déterminants relatifs, qui déterminent un nom (ex. en français : « un homme, *lequel* homme portait un chapeau »)

Différentes catégories détaillées étiquetant des déterminants relatifs et catégorie générale qui leur correspondent

Catégorie détaillée	Formes fléchies concernées	Catégorie générale
DetRelMS	masculin singulier	DetRel
DetRelFS	féminin singulier	DetRel
DetRelMP	masculin pluriel	DetRel
DetRelFP	féminin pluriel	DetRel

Différentes catégories détaillées étiquetant des pronoms relatifs et catégories générales qui leur correspondent

Catégorie détaillée	Formes fléchies concernées	Catégorie générale
PronRelMS	masculin singulier	PronRel
PronRelFS	féminin singulier	PronRel
PronRelMP	masculin pluriel	PronRel
PronRelFP	féminin pluriel	PronRel
PronRelNeut		PronRelNeut

5.1.4.8 Personnels

La catégorie générale des pronoms personnels dépend de leur fonction et de leur personne. Nous avons distingué :

- Les pronoms personnels sujets (ex. en français : « *je chante* », « *il va* »...), avec parmi eux :
 - PronPers1sSuj : Ceux de la première personne du singulier
 - PronPers2sSuj : Ceux de la deuxième personne du singulier
 - PronPers3sSuj : Ceux de la troisième personne du singulier
 - PronPers1pSuj : Ceux de la première personne du pluriel
 - PronPers2pSuj : Ceux de la deuxième personne du pluriel
 - PronPers3pSuj : Ceux de la troisième personne du pluriel
- Les pronoms personnels COD antéposés (ex. : « *lo* pausa », « *il me* regarde »...), avec parmi eux :
 - PronPers1sCodAp : Ceux de la première personne du singulier
 - PronPers2sCodAp : Ceux de la deuxième personne du singulier
 - PronPers3sCodAp : Ceux de la troisième personne du singulier
 - PronPers1pCodAp : Ceux de la première personne du pluriel
 - PronPers2pCodAp : Ceux de la deuxième personne du pluriel
 - PronPers3pCodAp : Ceux de la troisième personne du pluriel
- Les pronoms personnels COD postposés (ex. : « *pausa-lo* », « *regarde-moi* »...), avec parmi eux :
 - PronPers1sCodPp : Ceux de la première personne du singulier
 - PronPers2sCodPp : Ceux de la deuxième personne du singulier
 - PronPers3sCodPp : Ceux de la troisième personne du singulier
 - PronPers1pCodPp : Ceux de la première personne du pluriel
 - PronPers2pCodPp : Ceux de la deuxième personne du pluriel
 - PronPers3pCodPp : Ceux de la troisième personne du pluriel
- Les pronoms personnels COI antéposés (ex. : « *li* parla », « *il me* demande »...), avec parmi eux :
 - PronPers1sCoiAp : Ceux de la première personne du singulier
 - PronPers2sCoiAp : Ceux de la deuxième personne du singulier
 - PronPers3sCoiAp : Ceux de la troisième personne du singulier
 - PronPers1pCoiAp : Ceux de la première personne du pluriel
 - PronPers2pCoiAp : Ceux de la deuxième personne du pluriel
 - PronPers3pCoiAp : Ceux de la troisième personne du pluriel
- Les pronoms personnels COI postposés (ex. : « *parla-li* », « *demande-moi* »...), avec parmi eux :
 - PronPers1sCoiPp : Ceux de la première personne du singulier
 - PronPers2sCoiPp : Ceux de la deuxième personne du singulier
 - PronPers3sCoiPp : Ceux de la troisième personne du singulier
 - PronPers1pCoiPp : Ceux de la première personne du pluriel
 - PronPers2pCoiPp : Ceux de la deuxième personne du pluriel
 - PronPers3pCoiPp : Ceux de la troisième personne du pluriel
- Les pronoms personnels toniques (ex. : « *ne parli amb el* », « *el, es grand* », « *j'en parle avec lui* », « *moi, je suis grand* »...), avec parmi eux :
 - PronPers1sTon : Ceux de la première personne du singulier
 - PronPers2sTon : Ceux de la deuxième personne du singulier
 - PronPers3sTon : Ceux de la troisième personne du singulier
 - PronPers1pTon : Ceux de la première personne du pluriel
 - PronPers2pTon : Ceux de la deuxième personne du pluriel

- PronPers3pTon : Ceux de la troisième personne du pluriel
- Les pronoms personnels réfléchis et réciproques antéposés (ex. : « *me pencheni* », « *se baton* »...), avec parmi eux :
 - PronPers1sReflAp : Ceux de la première personne du singulier
 - PronPers2sReflAp : Ceux de la deuxième personne du singulier
 - PronPers3sReflAp : Ceux de la troisième personne du singulier
 - PronPers1pReflAp : Ceux de la première personne du pluriel
 - PronPers2pReflAp : Ceux de la deuxième personne du pluriel
 - PronPers3pReflAp : Ceux de la troisième personne du pluriel
- Les pronoms personnels réfléchis et réciproques antéposés (ex. : « *penchante* », « *saluez-vous* »...), avec parmi eux :
 - PronPers2sReflPp : Ceux de la deuxième personne du singulier
 - PronPers1pReflPp : Ceux de la première personne du pluriel
 - PronPers2pReflPp : Ceux de la deuxième personne du pluriel
- Les pronoms personnels réfléchis tonique (ex. : « *que'm vesteishi jo-medish* », « *se parla a el-meteis* »...), avec parmi eux :
 - PronPers1sReflTon : Ceux de la première personne du singulier
 - PronPers2sReflTon : Ceux de la deuxième personne du singulier
 - PronPers3sReflTon : Ceux de la troisième personne du singulier
 - PronPers1pReflTon : Ceux de la première personne du pluriel
 - PronPers2pReflTon : Ceux de la deuxième personne du pluriel
 - PronPers3pReflTon : Ceux de la troisième personne du pluriel
- Les pronoms personnels indéfinis et neutres, avec parmi eux :
 - PronPersIndefSuj : Le pronom personnel indéfini sujet (« *òm dit* »)
 - PronPersIndefReflTon : Les pronoms personnels indéfinis réfléchis tonique (« *far atencion a se* », « *ac guardar tà si-medish* », « *penser à soi* », « *le faire soi-même* »... en français « *soi* », « *soi-même* »)
 - PronPersNeutCodPp : Les pronoms personnels indéfinis COD postposés (« *digas-o* », « *répète-le* »)
 - PronPersNeutCoiAp : Les pronoms personnels indéfinis COI postposés (« *ne mangi* », « *j'en prends* »)
 - PronPersNeutCoiPp : Les pronoms personnels indéfinis COI postposés (« *manjane* », « *prends-en* »)

Catégories détaillées étiquetant des pronoms personnels et catégories générales qui leur correspondent

Catégorie détaillée	Formes fléchies concernées	Catégorie générale
PronPers1sSujM	masculin	PronPers1sSuj
PronPers1sSujF	féminin	PronPers1sSuj
PronPers1sCodApM	masculin	PronPers1sCodAp
PronPers1sCodApF	féminin	PronPers1sCodAp
PronPers1sCoiApM	masculin	PronPers1sCoiAp
PronPers1sCoiApF	féminin	PronPers1sCoiAp
PronPers1sCodPpM	masculin	PronPers1sCodPp

PronPers1sCodPpF	féminin	PronPers1sCodPp
PronPers1sCoiPpM	masculin	PronPers1sCoiPp
PronPers1sCoiPpF	féminin	PronPers1sCoiPp
PronPers1sTonM	masculin	PronPers1sTon
PronPers1sTonF	féminin	PronPers1sTon
PronPers1sReflApM	masculin	PronPers1sReflAp
PronPers1sReflApF	féminin	PronPers1sReflAp
PronPers1sReflPpM	masculin	PronPers1sReflPp
PronPers1sReflPpF	féminin	PronPers1sReflPp
PronPers1sReflTonM	masculin	PronPers1sReflTon
PronPers1sReflTonF	féminin	PronPers1sReflTon
PronPers2sSujM	masculin	PronPers2sSuj
PronPers2sSujF	féminin	PronPers2sSuj
PronPers2sCodApM	masculin	PronPers2sCodAp
PronPers2sCodApF	féminin	PronPers2sCodAp
PronPers2sCoiApM	masculin	PronPers2sCoiAp
PronPers2sCoiApF	féminin	PronPers2sCoiAp
PronPers2sCodPpM	masculin	PronPers2sCodPp
PronPers2sCodPpF	féminin	PronPers2sCodPp
PronPers2sCoiPpM	masculin	PronPers2sCoiPp
PronPers2sCoiPpF	féminin	PronPers2sCoiPp
PronPers2sTonM	masculin	PronPers2sTon
PronPers2sTonF	féminin	PronPers2sTon
PronPers2sReflApM	masculin	PronPers2sReflAp
PronPers2sReflApF	féminin	PronPers2sReflAp
PronPers2sReflPpM	masculin	PronPers2sReflPp
PronPers2sReflPpF	féminin	PronPers2sReflPp
PronPers2sReflTonM	masculin	PronPers2sReflTon
PronPers2sReflTonF	féminin	PronPers2sReflTon
PronPers3sSujM	masculin	PronPers3sSuj
PronPers3sSujF	féminin	PronPers3sSuj
PronPers3sCodApM	masculin	PronPers3sCodAp
PronPers3sCodApF	féminin	PronPers3sCodAp
PronPers3sCoiApM	masculin	PronPers3sCoiAp

PronPers3sCoiApF	féminin	PronPers3sCoiAp
PronPers3sCodPpM	masculin	PronPers3sCodPp
PronPers3sCodPpF	féminin	PronPers3sCodPp
PronPers3sCoiPpM	masculin	PronPers3sCoiPp
PronPers3sCoiPpF	féminin	PronPers3sCoiPp
PronPers3sTonM	masculin	PronPers3sTon
PronPers3sTonF	féminin	PronPers3sTon
PronPers3sReflApM	masculin	PronPers3sReflAp
PronPers3sReflApF	féminin	PronPers3sReflAp
PronPers3sReflPpM	masculin	PronPers3sReflPp
PronPers3sReflPpF	féminin	PronPers3sReflPp
PronPers3sReflTonM	masculin	PronPers3sReflTon
PronPers3sReflTonF	féminin	PronPers3sReflTon
PronPers1pSujM	masculin	PronPers1pSuj
PronPers1pSujF	féminin	PronPers1pSuj
PronPers1pCodApM	masculin	PronPers1pCodAp
PronPers1pCodvF	féminin	PronPers1pCodAp
PronPers1pCoiApM	masculin	PronPers1pCoiAp
PronPers1pCoiApF	féminin	PronPers1pCoiAp
PronPers1pCodPpM	masculin	PronPers1pCodPp
PronPers1pCodPpF	féminin	PronPers1pCodPp
PronPers1pCoiPpM	masculin	PronPers1pCoiPp
PronPers1pCoiPpF	féminin	PronPers1pCoiPp
PronPers1pTonM	masculin	PronPers1pTon
PronPers1pTonF	féminin	PronPers1pTon
PronPers1pReflApM	masculin	PronPers1pReflAp
PronPers1pReflApF	féminin	PronPers1pReflAp
PronPers1pReflPpM	masculin	PronPers1pReflPp
PronPers1pReflPpF	féminin	PronPers1pReflPp
PronPers1pReflTonM	masculin	PronPers1pReflTon
PronPers1pReflTonF	féminin	PronPers1pReflTon
PronPers2pSujM	masculin	PronPers2pSuj
PronPers2pSujF	féminin	PronPers2pSuj
PronPers2pCodApM	masculin	PronPers2pCodAp

PronPers2pCodApF	féminin	PronPers2pCodAp
PronPers2pCoiApM	masculin	PronPers2pCoiAp
PronPers2pCoiApF	féminin	PronPers2pCoiAp
PronPers2pCodPpM	masculin	PronPers2pCodPp
PronPers2pCodPpF	féminin	PronPers2pCodPp
PronPers2pCoiPpM	masculin	PronPers2pCoiPp
PronPers2pCoiPpF	féminin	PronPers2pCoiPp
PronPers2pTonM	masculin	PronPers2pTon
PronPers2pTonF	féminin	PronPers2pTon
PronPers2pReflApM	masculin	PronPers2pReflAp
PronPers2pReflApF	féminin	PronPers2pReflAp
PronPers2pReflPpM	masculin	PronPers2pReflPp
PronPers2pReflPpF	féminin	PronPers2pReflPp
PronPers2pReflTonM	masculin	PronPers2pReflTon
PronPers2pReflTonF	féminin	PronPers2pReflTon
PronPers3pSujM	masculin	PronPers3pSuj
PronPers3pSujF	féminin	PronPers3pSuj
PronPers3pCodApM	masculin	PronPers3pCodPp
PronPers3pCodApF	féminin	PronPers3pCodPp
PronPers3pCoiApM	masculin	PronPers3pCoiPp
PronPers3pCoiApF	féminin	PronPers3pCoiPp
PronPers3pCodPpM	masculin	PronPers3pCodAp
PronPers3pCodPpF	féminin	PronPers3pCodAp
PronPers3pCoiPpM	masculin	PronPers3pCoiAp
PronPers3pCoiPpF	féminin	PronPers3pCoiAp
PronPers3pTonM	masculin	PronPers3pTon
PronPers3pTonF	féminin	PronPers3pTon
PronPers3pReflApM	masculin	PronPers3pReflAp
PronPers3pReflApF	féminin	PronPers3pReflAp
PronPers3pReflPpM	masculin	PronPers3pReflPp
PronPers3pReflPpF	féminin	PronPers3pReflPp
PronPers3pReflTonM	masculin	PronPers3pReflTon
PronPers3pReflTonF	féminin	PronPers3pReflTon
PronPersIndefSuj		PronPersIndefSuj

PronPersIndefReflTon		PronPersIndefReflTon
PronPersNeutCodAp		PronPersNeutCodAp
PronPersNeutCodPp		PronPersNeutCodPp
PronPersNeutCoiAp		PronPersNeutCoiAp
PronPersNeutCoiPp		PronPersNeutCoiPp

5.1.4.9 Numéraux cardinaux

Parmi eux, on distingue :

- Les adjectifs, quand le numéral complète un nom qui est déjà déterminé (ex. : « los *dus ostaus* »), avec parmi eux :
 - **AdjNumCardS** : les adjectifs numéraux cardinaux singuliers (« zéro », « un », « una »...)
 - **AdjNumCardP** : les adjectifs numéraux cardinaux pluriels (« *dus* », « *tres* », « *vint* »...)
- Les déterminants numéraux, quand le numéral détermine un nom (ex. : « i a *dos ostals* »), avec parmi eux :
 - **DetNumCardS** : les déterminants numéraux cardinaux singuliers
 - **DetNumCardP** : les déterminants numéraux cardinaux pluriels
- Les pronoms numéraux, quand le numéral remplace un groupe nominal (ex. : « *demest los ostals, dos son blancs* »), avec parmi eux :
 - **PronNumCardS** : les pronoms numéraux cardinaux singuliers
 - **PronNumCardP** : les pronoms numéraux cardinaux pluriels
- Les noms numéraux, quand le numéral est utilisé pour se représenter lui-même (ex. : « *un parelh de dos* », « *escriu un quatre a la fin* », « *una candela en forma de tres* »), avec parmi eux :
 - **NomNumCardM** : nom numéral

Catégories détaillées étiquetant des adjectifs numéraux cardinaux et catégories générales qui leur correspondent

Catégorie détaillée	Formes fléchies concernées	Catégorie générale
AdjNumCardSM	masculin	AdjNumCardS
AdjNumCardSF	féminin	AdjNumCardS
AdjNumCardPM	masculin	AdjNumCardP
AdjNumCardPF	féminin	AdjNumCardP

Catégories détaillées étiquetant des déterminants numéraux cardinaux et catégories générales qui leur correspondent

Catégorie détaillée	Formes fléchies concernées	Catégorie générale
DetNumCardSM	masculin	DetNumCardS
DetNumCardSF	féminin	DetNumCardS
DetNumCardPM	masculin	DetNumCardP

Catégorie détaillée	Formes fléchies concernées	Catégorie générale
DetNumCardPF	féminin	DetNumCardP

Catégories détaillées étiquetant des pronoms numéraux cardinaux et catégories générales qui leur correspondent

Catégorie détaillée	Formes fléchies concernées	Catégorie générale
PronNumCardSM	masculin	PronNumCardS
PronNumCardSF	féminin	PronNumCardS
PronNumCardPM	masculin	PronNumCardP
PronNumCardPF	féminin	PronNumCardP

Catégories détaillées étiquetant des noms numéraux cardinaux et catégories générales qui leur correspondent

Catégorie détaillée	Formes fléchies concernées	Catégorie générale
NomNumCardMS	singulier	NomNumCardM
NomNumCardMP	pluriel	NomNumCardM

5.1.4.10 Numéraux ordinaux

Parmi eux, on distingue :

- AdjNumOrd : les adjectifs (ex. : « lo tresau còp », « es arribat detzen »).
- Les noms (ex. : « lo cinquen entrèt puèi », « qu'ei calculat au centau », « dobla sa tresena »...), avec parmi eux :
 - NomNumOrdM : les noms ordinaux masculins
 - NomNumOrdF : les noms ordinaux féminins

Catégories détaillées étiquetant des adjectifs numéraux ordinaux et catégorie générale qui leur correspond

Catégorie détaillée	Formes fléchies concernées	Catégorie générale
AdjNumOrdMS	masculin singulier	AdjNumOrd
AdjNumOrdFS	féminin singulier	AdjNumOrd
AdjNumOrdMP	masculin pluriel	AdjNumOrd
AdjNumOrdFP	féminin pluriel	AdjNumOrd

Catégories détaillées étiquetant des noms numéraux ordinaux et catégorie générale qui leur correspondent

Catégorie détaillée	Formes fléchies concernées	Catégorie générale
NomNumOrdMS	masculin singulier	NomNumOrdM
NomNumOrdMP	masculin pluriel	NomNumOrdM
NomNumOrdFS	féminin singulier	NomNumOrdF

Catégorie détaillée	Formes fléchies concernées	Catégorie générale
NomNumOrdFP	féminin pluriel	NomNumOrdF

5.1.4.11 Partitifs

L'article partitif (« de lait ») est étiqueté comme déterminant partitif **DetPart**.

Catégories détaillées étiquetant des articles partitifs et catégorie générale qui leur correspondent

Catégorie détaillée	Formes fléchies concernées	Catégorie générale
DetPart		DetPart

5.1.4.12 Cas des articles contractés

Pour étiqueter les articles contractés (« deu vilatge », « al primtemps », « jol pont », « dau pan », « a partir del camin », « deu son manteu »...), on utilise les catégories des mots qui les composent, séparées par +.

Différents étiquetages pour les articles contractés

Catégorie détaillée	Formes fléchies concernées	Catégorie générale
Prep+ArtDefMS	masculin singulier	Prep+ArtDef
Prep+ArtDefFS	féminin singulier	Prep+ArtDef
Prep+ArtDefMP	masculin pluriel	Prep+ArtDef
Prep+ArtDefFP	féminin pluriel	Prep+ArtDef
DetPart+ArtDefMS	masculin singulier	DetPart+ArtDef
DetPart+ArtDefFS	féminin singulier	DetPart+ArtDef
DetPart+ArtDefMP	masculin pluriel	DetPart+ArtDef
DetPart+ArtDefFP	féminin pluriel	DetPart+ArtDef

Les articles contractés peuvent parfois être ceux d'un nom propre (ex. : « tau Caire », « dels Estats-Units », « ara Marca »...). Dans ce cas, on étiquette comme suit :

Différents étiquetages pour les articles contractés avec des noms propres

Catégorie détaillée	Formes fléchies concernées	Catégorie générale
Prep+NomPropMS	masculin singulier	Prep+NomPropMS
Prep+NomPropFS	féminin singulier	Prep+NomPropFS
Prep+NomPropMP	masculin pluriel	Prep+NomPropMP
Prep+NomPropFP	féminin pluriel	Prep+NomPropFP

Différents étiquetages pour les articles contractés avec des déterminants possessifs

Catégorie détaillée	Formes fléchies concernées	Catégorie générale

Prep+DetPoss1sMS	masculin singulier	Prep+DetPoss1s
Prep+DetPoss1sFS	féminin singulier	Prep+DetPoss1s
Prep+DetPoss1sMP	masculin pluriel	Prep+DetPoss1s
Prep+DetPoss1sFP	féminin pluriel	Prep+DetPoss1s
Prep+DetPoss2sMS	masculin singulier	Prep+DetPoss2s
Prep+DetPoss2sFS	féminin singulier	Prep+DetPoss2s
Prep+DetPoss2sMP	masculin pluriel	Prep+DetPoss2s
Prep+DetPoss2sFP	féminin pluriel	Prep+DetPoss2s
Prep+DetPoss3sMS	masculin singulier	Prep+DetPoss3s
Prep+DetPoss3sFS	féminin singulier	Prep+DetPoss3s
Prep+DetPoss3sMP	masculin pluriel	Prep+DetPoss3s
Prep+DetPoss3sFP	féminin pluriel	Prep+DetPoss3s
Prep+DetPoss1pMS	masculin singulier	Prep+DetPoss1p
Prep+DetPoss1pFS	féminin singulier	Prep+DetPoss1p
Prep+DetPoss1pMP	masculin pluriel	Prep+DetPoss1p
Prep+DetPoss1pFP	féminin pluriel	Prep+DetPoss1p
Prep+DetPoss2pMS	masculin singulier	Prep+DetPoss2p
Prep+DetPoss2pFS	féminin singulier	Prep+DetPoss2p
Prep+DetPoss2pMP	masculin pluriel	Prep+DetPoss2p
Prep+DetPoss2pFP	féminin pluriel	Prep+DetPoss2p
Prep+DetPoss3pMS	masculin singulier	Prep+DetPoss3p
Prep+DetPoss3pFS	féminin singulier	Prep+DetPoss3p
Prep+DetPoss3pMP	masculin pluriel	Prep+DetPoss3p
Prep+DetPoss3pFP	féminin pluriel	Prep+DetPoss3p

Différents étiquetages pour les articles contractés avec des pronoms possessifs

Catégorie détaillée	Formes fléchies concernées	Catégorie générale
Prep+PronPoss1sMS	masculin singulier	Prep+PronPoss1s
Prep+PronPoss1sFS	féminin singulier	Prep+PronPoss1s
Prep+PronPoss1sMP	masculin pluriel	Prep+PronPoss1s
Prep+PronPoss1sFP	féminin pluriel	Prep+PronPoss1s
Prep+PronPoss2sMS	masculin singulier	Prep+PronPoss2s
Prep+PronPoss2sFS	féminin singulier	Prep+PronPoss2s
Prep+PronPoss2sMP	masculin pluriel	Prep+PronPoss2s

Prep+PronPoss2sFP	féminin pluriel	Prep+PronPoss2s
Prep+PronPoss3sMS	masculin singulier	Prep+PronPoss3s
Prep+PronPoss3sFS	féminin singulier	Prep+PronPoss3s
Prep+PronPoss3sMP	masculin pluriel	Prep+PronPoss3s
Prep+PronPoss3sFP	féminin pluriel	Prep+PronPoss3s
Prep+PronPoss1pMS	masculin singulier	Prep+PronPoss1p
Prep+PronPoss1pFS	féminin singulier	Prep+PronPoss1p
Prep+PronPoss1pMP	masculin pluriel	Prep+PronPoss1p
Prep+PronPoss1pFP	féminin pluriel	Prep+PronPoss1p
Prep+PronPoss2pMS	masculin singulier	Prep+PronPoss2p
Prep+PronPoss2pFS	féminin singulier	Prep+PronPoss2p
Prep+PronPoss2pMP	masculin pluriel	Prep+PronPoss2p
Prep+PronPoss2pFP	féminin pluriel	Prep+PronPoss2p
Prep+PronPoss3pMS	masculin singulier	Prep+PronPoss3p
Prep+PronPoss3pFS	féminin singulier	Prep+PronPoss3p
Prep+PronPoss3pMP	masculin pluriel	Prep+PronPoss3p
Prep+PronPoss3pFP	féminin pluriel	Prep+PronPoss3p

Attention, lorsqu'on lemmatise, on étiquette les lemmes de la même manière (ex. : « a+lo », « de+Lo Caire », « a partir de+los », « sus + los sons »...).

Les familles grammaticales sont notées de la même manière (ex. : « Prep+Pron », « Prep+NomProp », « Prep+Det »...).

5.1.5 Adverbes

Nous avons distingué :

- **Adv** : Les adverbes généraux (ex. : « *funciona normalament* », « *sovent*, son grises »...)
- **AdvComp** : Les adverbes comparatifs, qui s'associent avec « que » ou « com(a) » (ex. : « *mangi melhor que tu* », « *quei tan gran com tu* »...)
- **AdvConj** : Les adverbes conjonctifs, qui ont une fonction entre la conjonction et l'adverbe (ex. : « *çaquelà i vau* », « *i vau çaquelà* »...)
- **AdvExcl** : Les adverbes exclamatifs (ex. : « *coma as crescut !* », « *quant n'i a !* »...)
- **AdvInterr** : les adverbes interrogatifs (ex. : « *on vas ?* », « *quand arribas ?* »...)
- **AdvMod** : Les adverbes modaux, qui renseignent sur l'énonciateur, son sentiment, sa certitude (ex. : « *qu'i vau anar, solide !* », « *urosament*, siás vengut »...)
- **AdvNeg** : Les adverbes négatifs, les simples (ex. : « *i pensi pas* », « *e de moneda, brica !* », « *aquí tanpauc* »...) et les composés (ex. : « *n'i a pas brica* », « *ne volí pas tanpauc* »...)
- **AdvNumOrd** : Les adverbes ordinaux (« *primièrement* », « *segondament* »... « *primo* », « *segundo* »... « *en premier lieu* », « *en deuxième lieu* »...)

- **PreAdv** : les adverbes qui se placent avant un adverbe ou un adjectif et ne peuvent être utilisés seuls (ex. : « il fait très chaud », *« il y en a très. »). On y trouve « très », « si » et « plutôt » (quand il n'est pas AdvConj).
- **PronAdv** : les pronoms adverbiaux (ex. : « ne vòli », « qu'i vau »...)

Tous les adverbes font partie de la famille grammaticale **Adv**, sauf les adverbes conjonctifs qui font partie de la famille grammaticale **Conj**.

Catégories détaillées étiquetant des adverbes et catégories générales qui leur correspondent

Catégorie détaillée	Formes fléchies concernées	Catégorie générale
Adv		Adv
AdvComp		AdvComp
AdvConj		AdvConj
AdvExcl		AdvExcl
AdvInterr		AdvInterr
AdvMod		AdvMod
AdvNeg		AdvNeg
AdvNumOrd		AdvNumOrd
PreAdv		PreAdv
PronAdv		PronAdv

5.1.6 Conjonctions

Les conjonctions ont pour catégorie générale **ConjCoord** (conjonctions de coordination) et **ConjSub** (conjonctions de subordination).

Elles font partie de la famille grammaticale **Conj**.

Catégories détaillées étiquetant des conjonctions et catégorie générales qui leur correspondent

Catégorie détaillée	Formes fléchies concernées	Catégorie générale
ConjCoord		ConjCoord
ConjSub		ConjSub

5.1.7 Prépositions

Les prépositions ont pour catégorie générale **Prep**.

Elles font partie de la famille grammaticale **Prep**.

Catégorie détaillée étiquetant des prépositions et catégorie générale qui lui correspond

Catégorie détaillée	Formes fléchies concernées	Catégorie générale
Prep		Prep

5.1.8 Interjections

Les interjections ont pour catégorie générale **Interj**.

Elles font partie de la famille grammaticale **Interj**.

Catégorie détaillée étiquetant des interjections et catégorie générale qui lui correspond

Catégorie détaillée	Formes fléchies concernées	Catégorie générale
Interj		Interj

5.1.9 Particules

Parmi les particules, on a les étiquettes de catégorie générale :

- **PartNeg** (particule négative) pour le mot « ne » en français et dans les variétés de l'occitan où il existe et où ce n'est pas un énonciatif. Famille grammaticale : **PartNeg**.
- **Part** pour certaines particules comme en français « est-ce-que » Famille grammaticale : **Part**.
- **PartEnonc** pour les énonciatifs gascons. Famille grammaticale : **PartEnonc**.
 - Enonciatifs affirmatifs (« que me'n batí », « ja me'n vau »)
 - Enonciatifs conditionnels (« e m'ac avosse dit, que serí vienut »)
 - Enonciatifs exclamatifs (« be m'estonas ! »)
 - Enonciatifs d'incise (« "Atencion !", ce cridè », « "i vau", çò diguèt »)
 - Enonciatifs interrogatifs (« e fonciona ? »)
 - Enonciatifs négatifs (« ne pensi pas »)

Catégories détaillées étiquetant des particules et catégories générales qui leur correspondent

Catégorie détaillée	Formes fléchies concernées	Catégorie générale
Part		Part
PartNeg		PartNeg
PartEnonc		PartEnonc

5.1.10 URL

Les URL de sites et adresses mails, ainsi que les numéros de téléphone, ont la catégorie générale **Url**.

Famille grammaticale : **Url**.

Catégorie détaillée étiquetant des URL et catégorie générale qui lui correspond

Catégorie détaillée	Formes fléchies concernées	Catégorie générale
URL		URL

5.1.11 Ponctuations

Les signes de ponctuation on pour catégorie générale Ponct.

Famille grammaticale : Ponct.

Catégorie détaillée étiquetant des signes de ponctuation et catégorie générale qui lui correspond

Catégorie détaillée	Formes fléchies concernées	Catégorie générale
Ponct		Ponct

5.1.12 Catégorie grammaticale inconnue

Lorsque la catégorie d'un mot est inconnue, on peut utiliser la catégorie générale Inconnu. A celle-ci correspondent plusieurs catégories détaillées, selon si l'on connaît ou pas le genre et le nombre du mot.

On utilise également la famille grammaticale Inconnu.

Catégories détaillées étiquetant les mots inconnus et catégorie générale qui leur correspondent

Catégorie détaillée	Formes fléchies concernées	Catégorie générale
InconnuMS	masculin singulier	Inconnu
InconnuFS	féminin singulier	Inconnu
InconnuMP	masculin pluriel	Inconnu
InconnuFP	féminin pluriel	Inconnu
Inconnu		Inconnu

5.1.13 Catégories de famille grammaticale

Les étiquettes de famille grammaticale rassemblent des catégories générales ayant un fonctionnement grammatical similaire, pour aider à l'entraînement des machines (notamment pour l'étiquetage de corpus).

Tableau de correspondance catégories générales - catégories de famille grammaticale :

Catégorie générale	Catégorie de famille grammaticale
Adj	Adj
AdjDem	Adj
AdjIndef	Adj

Catégorie générale	Catégorie de famille grammaticale
AdjNumCardP	NumCard
AdjNumCardS	NumCard
AdjNumOrd	Adj
AdjPoss1p	Adj
AdjPoss1s	Adj
AdjPoss2p	Adj
AdjPoss2s	Adj
AdjPoss3p	Adj
AdjPoss3s	Adj
Adv	Adv
AdvComp	Adv
AdvConj	Adv
AdvExcl	Adv
AdvInterr	Adv
AdvMod	Adv
AdvNeg	Adv
AdvNumOrd	Adv
ConjCoord	Conj
ConjSub	Conj
DetDef	Det
DetDem	Det
DetExcl	Det
DetIndef	Det
DetIndefP	Det
DetIndefS	Det
DetInterr	Det
DetNumCardP	NumCard
DetNumCardS	NumCard
DetPart	DetPart
DetPart+DetDef	DetPart+DetDef
DetPoss1p	Det
DetPoss1s	Det
DetPoss2p	Det

Catégorie générale	Catégorie de famille grammaticale
DetPoss2s	Det
DetPoss3p	Det
DetPoss3s	Det
DetRel	Det
Inconnu	Inconnu
Interj	Interj
NomComF	NomCom
NomComFP	NomCom
NomComM	NomCom
NomComMP	NomCom
NomNumCardM	NumCard
NomNumOrdF	NomCom
NomNumOrdM	NomCom
NomPropFP	NomProp
NomPropFS	NomProp
NomPropMP	NomProp
NomPropMS	NomProp
NomProp	NomProp
PartEnonc	PartEnonc
PartNeg	PartNeg
Part	Part
Ponct	Ponct
PreAdv	Adv
Prep+DetDef	Prep+Det
Prep+DetPoss1p	Prep+Det
Prep+DetPoss1s	Prep+Det
Prep+DetPoss2p	Prep+Det
Prep+DetPoss2s	Prep+Det
Prep+DetPoss3p	Prep+Det
Prep+DetPoss3s	Prep+Det
Prep+NomPropFP	Prep+NomProp
Prep+NomPropFS	Prep+NomProp
Prep+NomPropMP	Prep+NomProp

Catégorie générale	Catégorie de famille grammaticale
Prep+NomPropMS	Prep+NomProp
Prep	Prep
Prep+PronPoss1p	Prep+PronPoss
Prep+PronPoss1s	Prep+PronPoss
Prep+PronPoss2p	Prep+PronPoss
Prep+PronPoss2s	Prep+PronPoss
Prep+PronPoss3p	Prep+PronPoss
Prep+PronPoss3s	Prep+PronPoss
PronAdv	PronAdv
PronDemNeut	Pron
PronDem	Pron
PronExclNeut	Pron
PronExcl	Pron
PronIndefNeut	Pron
PronIndefP	Pron
PronIndefS	Pron
PronInterrNeut	Pron
PronInterr	Pron
PronNumCardP	NumCard
PronNumCardS	NumCard
PronPers1pCodAp	PronPers
PronPers1pCodPp	PronPers
PronPers1pCoiAp	PronPers
PronPers1pCoiPp	PronPers
PronPers1pReflAp	PronPers
PronPers1pReflPp	PronPers
PronPers1pReflTon	PronPers
PronPers1pSuj	PronPers
PronPers1pTon	PronPers
PronPers1sCodAp	PronPers
PronPers1sCodPp	PronPers
PronPers1sCoiAp	PronPers
PronPers1sCoiPp	PronPers

Catégorie générale	Catégorie de famille grammaticale
PronPers1sReflAp	PronPers
PronPers1sReflPp	PronPers
PronPers1sReflTon	PronPers
PronPers1sSuj	PronPers
PronPers1sTon	PronPers
PronPers2pCodAp	PronPers
PronPers2pCodPp	PronPers
PronPers2pCoiAp	PronPers
PronPers2pCoiPp	PronPers
PronPers2pReflAp	PronPers
PronPers2pReflPp	PronPers
PronPers2pReflTon	PronPers
PronPers2pSuj	PronPers
PronPers2pTon	PronPers
PronPers2sCodAp	PronPers
PronPers2sCodPp	PronPers
PronPers2sCoiAp	PronPers
PronPers2sCoiPp	PronPers
PronPers2sReflAp	PronPers
PronPers2sReflPp	PronPers
PronPers2sReflTon	PronPers
PronPers2sSuj	PronPers
PronPers2sTon	PronPers
PronPers3pCodAp	PronPers
PronPers3pCodPp	PronPers
PronPers3pCoiAp	PronPers
PronPers3pCoiPp	PronPers
PronPers3pReflAp	PronPers
PronPers3pReflPp	PronPers
PronPers3pReflTon	PronPers
PronPers3pSuj	PronPers
PronPers3pTon	PronPers
PronPers3sCodAp	PronPers

Catégorie générale	Catégorie de famille grammaticale
PronPers3sCodPp	PronPers
PronPers3sCoiAp	PronPers
PronPers3sCoiPp	PronPers
PronPers3sReflAp	PronPers
PronPers3sReflPp	PronPers
PronPers3sReflTon	PronPers
PronPers3sSuj	PronPers
PronPers3sTon	PronPers
PronPersIndefReflTon	PronPers
PronPersIndefSuj	PronPers
PronPersNeutCodAp	PronPers
PronPersNeutCoiAp	PronPers
PronPersNeutCodPp	PronPers
PronPersNeutCoiPp	PronPers
PronPoss1p	Pron
PronPoss1s	Pron
PronPoss2p	Pron
PronPoss2s	Pron
PronPoss3p	Pron
PronPoss3s	Pron
PronRelNeut	Pron
PronRel	Pron
Url	Url
Verbe	Verbe

5.2 Utilisation de la norme eagles

Dans certains lexiques, les indications de catégories grammaticales habituelles peuvent être complétées par une indication qui suit la norme *Eagles* (<http://www.ilc.cnr.it/EAGLES96/annotate/node15.html>). Elle est indiquée entre les balises **<gram type="catEagles"></gram>**.

Elle est utilisée comme suit (les colonnes et valeurs grisées ne sont jamais utilisées, les attributs suivis d'une * ne peuvent pas avoir leur valeur à 0) :

5.2.1 Noms

Marqués N.

Attributs :

Ordre	Attribut	Valeurs				
		1	2	3	4	5
1	Type*	Commun	Propre			
2	Genre	Masculin	Féminin	Neutre		
3	Nombre	Singulier	Pluriel			
4	Cas	Nominatif	Génitif	Datif	Accusatif	Vocatif
5	Comptabilité	Comptable	Masse			

La valeur 0 est utilisée en cas d'attribut non pertinent ou dont la valeur est inconnue.

La comptabilité sert à différencier les mots comptables (comme bille, arbre...) des mots représentant des masses (lait, sable...).

Exemples :

- Nom commun masculin singulier masse : N11102
- Nom commun masculin pluriel comptable : N11201
- Nom commun féminin singulier (comptabilité inconnue) : N12100
- Nom commun féminin pluriel : N12101
- Nom propre masculin singulier : N21100

5.2.2 Verbes

Marqués V.

Attributs :

Ordre	Attribut	Valeurs			
		1 5	2 6	3 7	4 8
1	Personne	Première	Deuxième	Troisième	
2	Genre	Masculin	Féminin	Neutre	
3	Nombre	Singulier	Pluriel		
4	Finitude	Fini	Non fini		
5	Mode*	Indicatif	Subjonctif	Impératif	Conditionnel
		Infinitif	Participe	Gérondif	
6	Temps	Présent	Imparfait	Futur	Passé
7	Voix	Active	Passive		
8	Statut	Principal	Auxilliaire		
9	Aspect	Perfectif	Imperfectif		
10	Séparabilité	Non séparable	Séparable		
11	Réflexivité	Réfléchi	Non réfléchi		
12	Auxilliaire	Avoir	Être		
13	Fonction de l'auxilliaire	Primaire	Modal		
14	Type d'impératif	Affirmatif	Négatif		

La valeur 0 est utilisée en cas d'attribut non pertinent ou dont la valeur est inconnue.

Hormis à l'infinitif et au participe, il faut toujours indiquer la personne (1) et le nombre (3). Au participe passé, il faut indiquer le genre (2) et le nombre (3). A l'impératif, il faut indiquer le type (14).

Exemples :

- Verbe à l'infinitif : V000050000000000
- Verbe à la 1e pers du sg du présent de l'indicatif : V10101100000000
- Verbe pronominal à l'infinitif : V00005000001000
- Participe passé féminin singulier : V02106400000000
- Verbe à l'impératif affirmatif, 2e personne du pluriel : V20203100000001

5.2.3 Adjectifs

Marqués AJ.

Attributs :

Ordre	Attribut	Valeurs				
		1	2	3	4	5
1	Degré	Positif	Comparatif	Superlatif		
2	Genre*	Masculin	Féminin	Neutre		
3	Nombre*	Singulier	Pluriel			
4	Cas	Nominatif	Génitif	Datif	Accusatif	Vocatif

La valeur Ø est utilisée en cas d'attribut non pertinent ou dont la valeur est inconnue.

Exemples :

- Adjectif masculin singulier : AJ0110
- Adjectif masculin pluriel : AJ0120
- Adjectif féminin singulier : AJ0210
- Adjectif féminin pluriel : AJ0210

5.2.4 Pronoms et déterminants

Marqués PD.

Attention, les adjectifs démonstratifs, indéfinis, possessifs, interrogatifs, exclamatifs... sont considérés comme des déterminants.

Attributs :

Ordre	Attribut	Valeurs				
		1	2	3	4	5
1	Personne	Première	Deuxième	Troisième		
2	Genre	Masculin	Féminin	Neutre		
3	Nombre	Singulier	Pluriel			
4	Nombre de la personne (possessifs)	Singulier	Pluriel			
5	Cas	Nominatif	Génitif	Datif	Accusatif	Vocatif
6	Catégorie*	Pronom	Déterminant	Les deux		
7	Type de pronom	Démonstratif	Indéfini	Possessif	Interrogatif / Relatif	Personnel / Réfléchi
8	Type de déterminant	Démonstratif	Indéfini	Possessif	Interrogatif / Relatif	Partitif
9	Type pers./réfl. (précisions)	Personnel	Réfléchi	Réiproque		
10	Type int./rel. (précisions)	Interrogatif	Relatif	Exclamatif		
11	Poli	Poli	Familier			
12	Force	Faible	Fort			

La valeur 0 est utilisée en cas d'attribut non pertinent ou dont la valeur est inconnue.

L'attribut 4 (Nombre de la personne), sert uniquement pour les possessifs. Il indique le nombre complétant la personne (du singulier / du pluriel) par rapport au 3 (Nombre) qui concerne le nombre lié au contexte, à l'objet possédé. Par exemple, « mes » sera côté PD1021..., notre PD2012...

L'attribut 5 (Cas) sert à différencier les pronoms sujets (1. Nominatif), COD (4. Accusatif), COI (3. Datif), obliques (4. Oblique).

On n'utilise pas la valeur 6 (Les deux) pour l'attribut 6 (Catégorie). On met une étiquette pour le pronom, une étiquette pour le déterminant.

Il faut toujours que l'attribut 7 (type de pronom) ou l'attribut 8 (type de déterminant) soit rempli. Pour les pronoms et déterminants de type 4 (interrogatif/relatif), il faut remplir l'attribut 10 (type int./rel.).

L'attribut 11 (Poli) sert à distinguer par exemple le « vous » poli du « tu » familier pour la 2e personne du singulier.

L'attribut 12 (Force) sert à distinguer les pronoms conjoints (forts) des pronoms disjoints (faibles) et les pronoms toniques (forts) des pronoms atones (faibles).

Exemples :

- Prénom indéfini : PD0000001200000
- Prénom interrogatif masculin singulier : PD01101400100
- Adjectif possessif « tes » : PD212102030020, PD222102030020

5.2.5 Articles

Marqués AT.

Attributs :

Ordre	Attribut	Valeurs			
		1	2	3	4
1	Type*	Défini	Indéfini	Partitif	
2	Genre	Masculin	Féminin	Neutre	
3	Nombre	Singulier	Pluriel		
4	Cas	Nominatif	Génitif	Datif	Accusatif

La valeur 0 est utilisée en cas d'attribut non pertinent ou dont la valeur est inconnue.

Les attributs 2 (genre) et 3 (nombre) doivent être renseignés pour tous les articles sauf les partitifs.

Exemples :

- Article indéfini masculin singulier : AT2110
- Article défini féminin pluriel : AT1210
- Article partitif « de » : AT3000

5.2.6 Adverbes

Marqués AV.

Attributs :

Ordre	Attribut	Valeurs			
		1	2	3	4
1	Degré	Positif	Comparatif	Superlatif	Négatif
2	Type	Général	Degré		
3	Polarité	Wh-type	Non wh-type		
4	Wh-type	Interrogatif	Relatif	Exclamatif	

La valeur 0 est utilisée en cas d'attribut non pertinent ou dont la valeur est inconnue.

Les adverbes comparatifs sont melhor, mens... Les adverbes négatifs sont jamai, pus (généralement utilisés avec "pas").

Les adverbes de type « degré » sont ceux indiquant l'intensité, la quantité (tant, plan...).

Exemples :

- Adverbe : AV0000
- Adverbe interrogatif : AV0001

5.2.7 Adpositions

Marquées AP.

Attributs :

Ordre	Attribut	Valeurs	
		1	
1	Type*		Préposition

La valeur 0 est utilisée en cas d'attribut non pertinent ou dont la valeur est inconnue.

Exemples :

- Préposition : AP1

5.2.8 Conjonctions

Marquées C.

Attributs :

Ordre	Attribut	Valeurs	
		1	2
1	Type*	De coordination	De subordination

La valeur 0 est utilisée en cas d'attribut non pertinent ou dont la valeur est inconnue.

Exemples :

- Conjonction de coordination : C1
- Conjonction de subordination : C2

5.2.9 Numéraux

Marqués NU.

Attributs :

Ordre	Attribut	Valeurs			
		1	2	3	4
1	Type*	Cardinal	Ordinal		
2	Genre	Masculin	Féminin	Neutre	
3	Nombre*	Singulier	Pluriel		
4	Cas	Nominatif	Génitif	Datif	Accusatif
5	Fonction*	Pronom	Déterminant	Adjectif	

La valeur 0 est utilisée en cas d'attribut non pertinent ou dont la valeur est inconnue.

Les adjectifs numéraux cardinaux sont ceux employés comme suit « los dus amics ». Les pronoms numéraux cardinaux sont ceux employés comme suit « i a d'ostals, dos son barrats ». Les déterminants cardinaux sont ceux employés comme suit « dos amics ». Les « noms numéraux » (un parelh de dos, un cinquen del pastís...) sont catégorisés comme des noms, pas comme des numéraux.

Exemples :

- Adjectif numéral ordinal masculin singulier : NU21103
- Adjectif numéral cardinal : NU10203

5.2.10 Interjections

Marquées I.

Pas d'attribut.

5.2.11 Unique

Marqués U.

Utilisés pour les catégories avec un seul ou un très petit nombre de membres, comme les particules négatives, qui ne peuvent être assignées à une autre des catégories précédentes.

Attributs :

Ordre	Attribut	Valeurs			
		1 5	2 6	3 7	4 8
1	Type*		Particule négative		Seconde particule négative
					Enonciatif
2	Type d'énonciatif	Affirmatif	Négatif	Interrogatif	Exclamatif
		Incise	Condition		

La valeur 0 est utilisée en cas d'attribut non pertinent ou dont la valeur est inconnue.

En occitan, le type 2 (particule négative) est utilisé pour « pas ». En français, où il y a deux particules négatives, on utilise le type 2 (particule négative) pour « ne » et le type 4 (seconde particule négative) pour « pas ».

Le type d'énonciatif est obligatoire si le type est 8 (énonciatif).

5.2.12 Résiduel

Marqués R.

Pour les éléments ne rentrant dans aucune catégorie.

Attributs :

Ordre	Attribut	Valeurs		
		1 4	2 5	3 6
1	Type*	Mot étranger	Formule	Symbole
		Acronyme	Abréviation	Non classifié
2	Nombre	Singulier	Pluriel	
3	Genre	Masculin	Féminin	Neutre

La valeur 0 est utilisée en cas d'attribut non pertinent ou dont la valeur est inconnue.

La plupart des sigles sont cotés comme des noms ou des adjectifs (CD, BCBG...) et les abréviations, autant que possible, sont cotées comme le mot complet qu'elles abrègent.

Exemples :

- suffixe féminin singulier : R612

5.2.13 Ponctuations

Marquées PU.

5.3 Utilisation de la norme Grace

Dans certains lexiques, les indications de catégories grammaticales habituelles peuvent être complétées par une indication qui suit la norme Grace. Elle est indiquée entre les balises **<gram type="catGrace"></gram>**.

L'utilisation de la norme Grace pour l'occitan suit les préconisations mises en place par CLLE-ERSS dans le cadre de la création du lexique de formes fléchies Loflòc.

6 Liste des valeurs possibles pour les attributs « norm »

6.1 Pour la balise `<case>`

- anteposed COD
- postposed COD
- anteposed COI
- postposed COI
- anteposed reflexive
- postposed reflexive
- tonic reflexive
- subject
- tonic

Cette liste est complète.

6.2 Pour la balise `<gen>`

- feminine
- masculine
- neuter

Cette liste est complète.

6.3 Pour la balise `<gram type="polarity">`

- affirmative
- negative

Cette liste est complète.

6.4 Pour la balise `<iType type="accord">`

- gender variable
- invariable
- number variable
- optional agreement
- variable

Cette liste peut être complétée.

6.5 Pour la balise `<iType type="altvb">`

- e/e
- è/e
- é/e
- iè/e
- ò/o
- ó/o
- o/o
- òu/ou

- uè/o
- ue/u
- uè/u
- ue/u

Cette liste peut être complétée.

6.6 Pour la balise `<iType type="inchoative">`

- inchoative
- not inchoative

Cette liste est complète.

6.7 Pour la balise `<lbl type="nature">` (sous `<form>`)

- affix
- abbreviation
- contraction
- full
- international unit
- prefix
- sigle
- suffix
- symbol
- synonym
- variant

Cette liste est complète.

6.8 Pour la balise `<lbl type="reference">` (sous `<xr>`)

- abbreviations
- antonyms
- augmentatives
- binName
- cf
- cognates
- contractions
- demonyms
- deriveds
- diminutives
- full
- hyperonyms
- keywords
- lemma
- sigle
- symbol
- synonyms
- variants

Cette liste peut être complétée.

6.9 Pour la balise <mood>

- conditional
- gerundive
- imperative
- indicative
- infinitive
- participle
- subjunctive

Cette liste peut être complétée.

6.10 Pour la balise <number>

- plural
- singular

Cette liste est complète.

6.11 Pour la balise <per>

- 1
- 2
- 3

Cette liste est complète.

6.12 Pour la balise <pos>

- adjective
- adverb
- article
- conjunction
- determiner
- exclamation
- interjection
- locution
- mimologism
- noun
- numeral adjective
- onomatopeia
- prefix
- preposition
- pronoun
- suffix
- symbol
- verb

Cette liste peut être complétée.

6.13 Pour la balise `<subc>`

Pour les adverbes :

- manner
- place
- time

Pour les verbes :

- indirect intransitive
- indirect transitive
- intransitive
- pronominal
- transitive

Cette liste peut être complétée.

6.14 Pour la balise `<tns>`

- future
- imperfect
- past
- perfect
- present

Cette liste peut être complétée.

6.15 Pour la balise `<usg type="dom">`

- accounting
- administration
- aeronautics
- agriculture
- agronomics
- alchemy
- anatomy
- anthropology
- apiculture
- arboriculture
- archeology
- architecture
- armory
- art
- astrology
- astronautics
- astronomy
- audiovisual
- automatism
- automobile
- aviation
- biochemistry

- biology
- botany
- building
- business
- butchery
- cartography
- chemistry
- cinema
- circus
- cooking
- corrida
- couture
- dance
- didactics
- drama
- economics
- electricity
- electronics
- entomology
- ethology
- fishing
- game
- genetics
- geography
- geology
- geometry
- geophysics
- grammar
- heraldry
- history
- hunting
- ichthyology
- industry
- informatics
- law
- linguistics
- literature
- mathematics
- mechanics
- medicine
- metalworking
- meteorology
- metrology
- military
- mineralogy
- music
- mycology

- mythology
- numismatics
- nutrition
- optical
- ornithology
- painting
- paleography
- paleontology
- pharmacy
- philosophy
- phonetics
- photography
- physics
- politics
- printing
- psychology
- radio
- religion
- rhetoric
- sailing
- school
- science
- sculpture
- sport
- statistics
- surgery
- technology
- toponymy
- zoology

Cette liste peut être complétée.

6.16 Pour la balise `<usg type="plev">`

- absolutely (absolument, sens absolu)
- every meaning (tous les sens)
- not attested (pas attesté)
- particularly (en particulier, sens particulier)
- rare (peu usité, rare)
- strong sense (sens fort)
- used (usité, très usité, usuel, attesté par l'usage...)
- usually (sens général, souvent, usuel)

Cette liste peut être complétée.

6.17 Pour la balise `<usg type="register">`

- affectionate (affectueux)
- argot
- burlesque

- *childish* (langage enfantin)
- *classical* (langue classique)
- *crude* (trivial, vulgaire)
- *familiar* (familier)
- *formal* (soutenu)
- *improper* (impropre)
- *ironic* (ironique)
- *literary* (littéraire)
- *not pejorative* (non péjoratif)
- *oral* (langue parlée)
- *pejorative* (péjoratif)
- *pleasant* (plaisant)
- *poetic* (poétique)
- *popular* (populaire)
- *scientific* (scientifique)
- *school* (scolaire)
- *technical* (technique)

Cette liste peut être complétée.

6.18 Pour la balise `<usg type="style">`

- *abstract* (abstrait)
- *concrete* (concret, sens matériel)
- *figurative* (sens figuré)
- *literal* (littéral, sens propre)
- *usual* (langue usuelle)

Cette liste peut être complétée.

6.19 Pour la balise `<usg type="time">`

- *archaic* (archaïque)
- *medieval* (Moyen-Âge)
- *neologism* (néologisme)
- *now* (aujourd'hui)
- *old* (vieux, ancien, vieilli)
- *traditional* (traditionnel)
- *s. XVIII* (18e siècle)

Cette liste peut être complétée.

6.20 Pour les indications de catégories grammaticales

6.20.1 Contenu des labels des catégories détaillées

Voir le fichier *categories_grammaticales_congres.xlsx* (feuille « *catdet* »), qui a pour colonnes :

- L'étiquette unique de catégorie (valeur mise dans l'attribut « **norm** » du `<gram type="catDet">`)

- Les informations morphologiques en occitan (valeur mise dans l'attribut « **norm** » du **<lbl type="catDet" xml:lang="oc-grclass">** si la forme est après son lemme)
- Les informations morphologiques en français (valeur mise dans l'attribut « **norm** » du **<lbl type="catDet" xml:lang="fr">** si la forme est après son lemme)
- Le nom complet de la catégorie en occitan (valeur mise dans l'attribut « **norm** » du **<lbl type="catDet" xml:lang="oc-grclass">** si la forme est sans son lemme)
- Le nom complet de la catégorie en français (valeur mise dans l'attribut « **norm** » du **<lbl type="catDet" xml:lang="fr">** si la forme est sans son lemme)
- Les informations morphologiques abrégées en occitan (valeur entre balises du **<lbl type="catDet" xml:lang="oc-grclass">** si la forme est après son lemme)
- Les informations morphologiques abrégées en français (valeur entre balises du **<lbl type="catDet" xml:lang="fr">** si la forme est après son lemme)
- Le nom abrégé de la catégorie en occitan (valeur entre balises du **<lbl type="catDet" xml:lang="oc-grclass">** si la forme est sans son lemme)
- Le nom abrégé de la catégorie en français (valeur entre balises du **<lbl type="catDet" xml:lang="fr">** si la forme est sans son lemme)
- L'étiquette de la catégorie générale correspondante (valeur mise dans l'attribut « **norm** » du **<gram type="catGen">**)
- L'étiquette de la catégorie de famille grammaticale correspondante (valeur mise dans l'attribut « **norm** » du **<gram type="catFam">**)

6.20.2 Contenu des labels des catégories générales

Voir le fichier categories_grammaticales_congres.xlsx (feuille « catgen »), qui a pour colonnes :

- L'étiquette unique de catégorie (valeur mise dans l'attribut « **norm** » du **<gram type="catGen">**)
- Le nom complet de la catégorie en occitan (valeur mise dans l'attribut « **norm** » du **<lbl type="catGen" xml:lang="oc-grclass">**)
- Le nom complet de la catégorie en français (valeur mise dans l'attribut « **norm** » du **<lbl type="catGen" xml:lang="fr">**)
- Le nom abrégé de la catégorie en occitan (valeur entre les balises du **<lbl type="catGen" xml:lang="oc-grclass">**)
- Le nom abrégé de la catégorie en français (valeur entre les balises du **<lbl type="catGen" xml:lang="fr">**)
- L'étiquette de la catégorie de famille grammaticale correspondante (valeur mise dans l'attribut « **norm** » du **<gram type="catFam">**)

6.20.3 Etiquettes des différents pronoms personnels

Lg	Pronom	catGen	catDet	Exemple
fr	elle	PronPers3sSujF PronPers3sTonF	PronPers3sSuj PronPers3sTon	<i>elle</i> parle parler <i>d'elle</i> ; <i>elle</i> , elle parle
fr	elle-même	PronPers3sReflTonF	PronPers3sReflTon	<i>elle</i> est <i>elle-même</i> grande
fr	elles	PronPers3pSujF PronPers3pTonF	PronPers3pSuj PronPers3pTon	<i>elles</i> parlent parler <i>d'elles</i> ; <i>elles</i> , elles sont
fr	elles-	PronPers3pReflTonF	PronPers3pReflTon	étant <i>elles-mêmes</i> grandes

	mêmes			
fr	en	PronPersNeutCodAp PronPersNeutCodPp PronPersNeutCoiAp PronPersNeutCoiPp	PronPersNeutCodAp PronPersNeutCodPp PronPersNeutCoiAp PronPersNeutCoiPp	j'en veux regardes-en j'en parle parles-en
fr	eux	PronPers3pTonM	PronPers3pTon	parler d'eux ; eux, ils sont
fr	eux-mêmes	PronPers3pReflTonM	PronPers3pReflTon	étant eux-mêmes grands
fr	il	PronPers3sSujM	PronPers3sSuj	il parle
fr	ils	PronPers3pSujM	PronPers3pSuj	ils parlent
fr	je	PronPers1sSujM PronPers1sSujF	PronPers1sSuj	je suis beau je suis belle
fr	l'	PronPers3sCodApM PronPers3sCodApF	PronPers3sCodAp	de l'appeler Jean de l'appeler Jeanne
fr	la	PronPers3sCodApF PronPers3sCodPpF	PronPers3sCodAp PronPers3sCodPp	je la vois regarde-la
fr	le	PronPers3sCodApM PronPers3sCodPpM	PronPers3sCodAp PronPers3sCodPp	je le vois regarde-le
fr	les	PronPers3pCodApM PronPers3pCodApF PronPers3pCodPpM PronPers3pCodPpF	PronPers3pCodAp PronPers3pCodPp	je les vois à eux je les vois à elles regarde-les à eux regarde-les à elles
fr	leur	PronPers3pCoiApM PronPers3pCoiApF PronPers3pCoiPpM PronPers3pCoiPpF	PronPers3pCoiAp PronPers3pCoiPp	je leur parle à eux je leur parle à elles parle-leur à eux parle-leur à elles
fr	lui	PronPers3sCoiApM PronPers3sCoiApF PronPers3sCoiPpM PronPers3sCoiPpF PronPers3sTonM	PronPers3sCoiAp PronPers3sCoiPp PronPers3sTon	je lui parle à lui je lui parle à elle parle-lui à lui parle-lui à elle parler de lui ; lui, il parle
fr	lui-même	PronPers3sReflTonM	PronPers3sReflTon	il est lui-même grand
fr	me	PronPers1sCodApM PronPers1sCodApF PronPers1sCoiApM PronPers1sCoiApF PronPers1sReflApM PronPers1sReflApF	PronPers1sCodAp PronPers1sCoiAp PronPers1sReflAp	il me traite d'idiot il me traite d'itiote il me parle (homme) il me parle (femme) je me peigne (homme) je me peigne (femme)
fr	m'	PronPers1sCodApM PronPers1sCodApF PronPers1sCoiApM PronPers1sCoiApF PronPers1sReflApM PronPers1sReflApF	PronPers1sCodAp PronPers1sCoiAp PronPers1sReflAp	il m'apelle « idiot » il m'apelle « idiote » il m'écoute (homme) il m'écoute (femme) je m'étais peigné (homme) je m'étais peigné (femme)
fr	moi	PronPers1sTonM PronPers1sTonF	PronPers1sTon	moi, je suis beau moi, je suis belle

		PronPers1sCodPpM PronPers1sCodPpF PronPers1sCoiPpM PronPers1sCoiPpF	PronPers1sCodPp PronPers1sCoiPp	traite- <i>moi</i> d'idiot traite- <i>moi</i> d'idiote parle- <i>moi</i> (homme) parle- <i>moi</i> (femme)
fr	moi-même	PronPers1sReflTonM PronPers1sReflTonF	PronPers1sReflTon	je suis <i>moi-même</i> grand je suis <i>moi-même</i> grande
fr	nous	PronPers1pCodApM PronPers1pCodApF PronPers1pCodPpM PronPers1pCodPpF PronPers1pCoiApM PronPers1pCoiApF PronPers1pCoiPpM PronPers1pCoiPpF PronPers1pReflApM PronPers1pReflApF PronPers1pReflPpM PronPers1pReflPpF PronPers1pSujM PronPers1pSujF PronPers1pTonM PronPers1pTonF	PronPers1pCodAp PronPers1pCodPp PronPers1pCoiAp PronPers1pCoiPp PronPers1pReflAp PronPers1pReflPp PronPers1pSuj PronPers1pTon	tu <i>nous</i> trouves beaux tu <i>nous</i> trouves belles traite- <i>nous</i> d'idiots traite- <i>nous</i> d'idiotes il <i>nous</i> parle (hommes) il <i>nous</i> parle (femmes) parle- <i>nous</i> (hommes) parle- <i>nous</i> (femmes) nous <i>nous</i> peignons (h.) nous <i>nous</i> peignons (f.) peignons- <i>nous</i> (hommes) peignons- <i>nous</i> (femmes) <i>nous</i> sommes beaux <i>nous</i> sommes belles <i>nous</i> , <i>nous</i> parlons ; elle parle de <i>nous</i> (hommes) <i>nous</i> , <i>nous</i> parlons ; elle parle de <i>nous</i> (femmes)
		PronPers1pReflTonM PronPers1pReflTonF	PronPers1pReflTon	étant <i>nous-mêmes</i> grands étant <i>nous-mêmes</i> grandes
		PronPersIndefSuj	PronPersIndefSuj	<i>on</i> dit que
		PronPers3sReflApM PronPers3sReflApF PronPers3pReflApM PronPers3pReflApF	PronPers3sReflAp PronPers3pReflAp	il <i>se</i> peigne elle <i>se</i> peigne ils <i>se</i> peignent elles <i>se</i> peignent
		PronPers3sReflApM PronPers3sReflApF PronPers3pReflApM PronPers3pReflApF	PronPers3sReflAp PronPers3pReflAp	il <i>s'était</i> peigné elle <i>s'était</i> peigné ils <i>s'étaient</i> peignés elles <i>s'étaient</i> peignées
	soi	PronPersIndefReflTon	PronPersIndefReflTon	<i>on</i> pense à <i>soi</i>
	soi-même	PronPersIndefReflTon	PronPersIndefReflTon	penser à <i>soi-même</i>
	PronPers2sCodApM PronPers2sCodApF PronPers2sCoiApM PronPers2sCoiApF PronPers2sReflApM PronPers2sReflApF	PronPers2sCodAp PronPers2sCoiAp PronPers2sReflAp	je <i>te</i> trouve beau je <i>te</i> trouve belle je <i>te</i> parle (homme) je <i>te</i> parle (femme) tu <i>te</i> peignes (homme) tu <i>te</i> peignes (femme)	
	PronPers2sCodApM PronPers2sCodApF PronPers2sCoiApM PronPers2sCoiApF PronPers2sReflApM	PronPers2sCodAp PronPers2sCoiAp PronPers2sReflAp	il <i>t'apelle</i> « idiot » il <i>t'apelle</i> « idiote » il <i>t'écoute</i> (homme) il <i>t'écoute</i> (femme) tu <i>t'étais</i> peigné (homme)	

		PronPers2sRefApF		tu t'étais peigné (femme)
fr	toi	PronPers3pRefPpM PronPers3pRefPpF PronPers2sTonM PronPers2sTonF	PronPers3pRefPp PronPers2sTon	peigne-toi (homme) peigne-toi (femme) toi, tu parles ; il parle de toi (homme) toi, tu parles ; il parle de toi (femme)
fr	toi-même	PronPers2sRefTonM PronPers2sRefTonF	PronPers2sRefTon	tu es toi-même grand tu es toi-même grande
fr	tu	PronPers2sSujM PronPers2sSujF	PronPers2sSuj	tu es beau tu es belle
fr	vous	PronPers2pCodApM PronPers2pCodApF PronPers2pCodPpM PronPers2pCodPpF PronPers2pCoiApM PronPers2pCoiApF PronPers2pCoiPpM PronPers2pCoiPpF PronPers2pRefApM PronPers2pRefApF PronPers2pRefPpM PronPers2pRefPpF PronPers2pSujM PronPers2pSujF PronPers2pTonM PronPers2pTonF	PronPers2pCodAp PronPers2pCodPp PronPers2pCoiAp PronPers2pCoiPp PronPers2pRefAp PronPers2pRefPp PronPers2pSuj PronPers2pTon	il vous trouve beaux il vous trouve belles traite-vous d'idiots traite-vous d'idiotes il vous parle (homme) il vous parle (femme) parle-vous (hommes) parle-vous (femmes) vous vous peignez (h.) vous vous peignez (f.) peignez-vous (hommes) peignez-vous (femmes) vous êtes beaux vous êtes belles vous, vous parlez ; je parle de vous (hommes) vous, vous parlez ; je parle de vous (femmes)
fr	vous-mêmes	PronPers2pRefTonM PronPers2pRefTonF	PronPers2pRefTon	étant vous-mêmes grands étant vous-mêmes grandes
fr	vous-même	PronPers2pRefTonM PronPers2pRefTonF	PronPers2pRefTon	étant vous-mêmes grand étant vous-mêmes grande
fr	y	PronPersNeutCoiAp PronPersNeutCoiPp	PronPersNeutCoiAp PronPersNeutCoiPp	tu y manges manges-y
oc	ac / o	PronPersNeutCodAp PronPersNeutCodPp	PronPersNeutCodAp PronPersNeutCodPp	o dises ; qu'ac dises digas-o ; diga'l'ac
oc	'c (gasc)	PronPersNeutCodPp	PronPersNeutCodPp	diga'c
oc	el / eth / eu	PronPers3sTonM	PronPers3sTon	el, parla ; parli d'el
oc	el-meiteis...	PronPers3sRefTonM	PronPers3sRefTon	o fa el-meteis
oc	ela / era	PronPers3sTonF	PronPers3sTon	ela, parla ; parli d'ela
oc	ela-meiteissa...	PronPers3sRefTonF	PronPers3sRefTon	o fa ela-meteissa
oc	elas / eras	PronPers3pTonF	PronPers3pTon	elas, parlan ; parli d'elas
oc	el(e)s / eths	PronPers3pTonM	PronPers3pTon	eles, parlan ; parli d'eles

oc	eles-meteisses...	PronPers3sReflTonM	PronPers3sReflTon	o fan <i>eles-meteisses</i>
oc	en / ne	PronPersNeutCodAp PronPersNeutCodPp PronPersNeutCoiAp PronPersNeutCoiPp	PronPersNeutCodAp PronPersNeutCodPp PronPersNeutCoiAp PronPersNeutCoiPp	<i>ne</i> vòli gaita- <i>ne</i> <i>ne</i> parli parla- <i>ne</i>
oc	n'	PronPersNeutCodAp PronPersNeutCodPp PronPersNeutCoiAp PronPersNeutCoiPp	PronPersNeutCodAp PronPersNeutCodPp PronPersNeutCoiAp PronPersNeutCoiPp	<i>n'apreni</i> minja- <i>n'</i> <i>n'expliqui</i> explica- <i>n'</i>
oc	'n	PronPersNeutCodAp PronPersNeutCodPp (gasc) PronPersNeutCoiAp PronPersNeutCoiPp (gasc)	PronPersNeutCodAp PronPersNeutCodPp PronPersNeutCoiAp PronPersNeutCoiPp	de' <i>n</i> minjar minja' <i>n</i> de' <i>n</i> parlar parla' <i>n</i>
oc	i	PronPersNeutCoiAp PronPersNeutCoiPp	PronPersNeutCoiAp PronPersNeutCoiPp	<i>i</i> mangi manjatz- <i>i</i>
oc	'i (gasc)	PronPersNeutCoiAp PronPersNeutCoiPp	PronPersNeutCoiAp PronPersNeutCoiPp	de' <i>i</i> minjar minjà' <i>i</i>
oc	ieu / jo	PronPers1sTonM PronPers1sTonF	PronPers1sTon	<i>ieu</i> , soi bèl ; parla de <i>ieu</i> <i>ieu</i> , soi bèla ; parla de <i>ieu</i>
oc	ieu-meteis...	PronPers1sReflTonM	PronPers1sReflTon	o fau <i>ieu-meteis</i>
oc	ieu-meteissa...	PronPers1sReflTonF	PronPers1sReflTon	o fau <i>ieu-meteissa</i>
oc	l'	PronPers3sCodApM PronPers3sCodApF PronPers3sCodPpM PronPers3sCodPpF PronPers3sCoiApM (gasc) PronPers3sCoiApF (gasc) PronPers3sCoiPpM (gasc) PronPers3sCoiPpF (gasc)	PronPers3sCodAp PronPers3sCodPp PronPers3sCoiAp (gasc) PronPers3sCoiPp (gasc)	<i>l'apèli</i> <i>l'apèli</i> minja- <i>l'i</i> minja- <i>l'i</i> que <i>l'expliqui</i> (òme) que <i>l'expliqui</i> (femna) retròba- <i>l'i</i> (òme) retròba- <i>l'i</i> (femna)
oc	la	PronPers3sCodApF PronPers3sCodPpF PronPers3sCoiApF (gasc) PronPers3sCoiPpF (gasc)	PronPers3sCodAp PronPers3sCodPp PronPers3sCoiAp (gasc) PronPers3sCoiPp (gasc)	<i>la</i> gaiti gaita- <i>la</i> <i>la</i> parli parla- <i>la</i>
oc	las	PronPers3pCodFAp PronPers3pCodFPp PronPers3pCoiFAp (gasc) PronPers3pCoiFPp (gasc)	PronPers3pCodAp PronPers3pCodPp PronPers3pCoiAp (gasc) PronPers3pCoiPp (gasc)	<i>las</i> gaiti gaita- <i>las</i> <i>las</i> parli parla- <i>las</i>
oc	lei(s)	PronPers3pCodApM PronPers3pCodApF	PronPers3pCodApF	<i>lei</i> trapi bèus <i>lei</i> trapi bèlas

		PronPers3pCodPpM PronPers3pCodPpF	PronPers3pCodPp	regarda- <i>lei</i> (òmes) regarda- <i>lei</i> (femnas)
oc	li	PronPers3sCoiApM PronPers3sCoiApF PronPers3sCoiPpM PronPers3sCoiPpF PronPers3pCoiApM (prov) PronPers3pCoiApF (prov) PronPers3pCoiPpM (prov) PronPers3pCoiPpF (prov)	PronPers3sCoiAp PronPers3sCoiPp PronPers3pCoiAp (prov) PronPers3pCoiPp (prov)	<i>li</i> parli a el <i>li</i> parli a ela parla- <i>li</i> a el parla- <i>li</i> a ela <i>li</i> parli a elei (òmes) <i>li</i> parli a elei (femnas) parla- <i>li</i> a elei (òmes) parla- <i>li</i> a elei (femnas)
oc	lo	PronPers3sCodApM PronPers3sCodPpM PronPers3sCoiApM (gasc) PronPers3sCoiPpM (gasc)	PronPers3sCodAp PronPers3sCodPp PronPers3sCoiAp (gasc) PronPers3sCoiAp (gasc)	<i>lo</i> gaiti gaita- <i>lo</i> <i>lo</i> parli parla- <i>lo</i>
oc	'u (gasc)	PronPers3sCodApM PronPers3sCodApF PronPers3sCodPpM PronPers3sCodPpF PronPers3sCoiApM (gasc) PronPers3sCoiApF (gasc) PronPers3sCoiPpM (gasc) PronPers3sCoiPpF (gasc)	PronPers3sCodAp PronPers3sCodPp PronPers3sCoiAp (gasc) PronPers3sCoiAp (gasc)	de' <i>u</i> minjar (òme) de' <i>u</i> minjar (femna) minja' <i>u</i> (òme) minja' <i>u</i> (femna) de' <i>u</i> parlar (òme) de' <i>u</i> parlar (femna) parla' <i>u</i> (òme) parla' <i>u</i> (femna)
oc	lor	PronPers3pCoiApM PronPers3pCoiApF PronPers3pCoiPpM PronPers3pCoiPpF	PronPers3pCoiAp PronPers3pCoiPp	<i>lor</i> parli a els <i>lor</i> parli a elas parla- <i>lor</i> a els parla- <i>lor</i> a elas
oc	'us (gasc)	PronPers3pCodApM PronPers3pCodApF PronPers3pCodPpM PronPers3pCodPpF PronPers3pCoiApM (gasc) PronPers3pCoiApF (gasc) PronPers3pCoiPpF (gasc) PronPers3pCoiPpM (gasc)	PronPers3pCodAp PronPers3pCodPp PronPers3pCoiAp (gasc) PronPers3pCoiAp (gasc)	de' <i>us</i> minjar (òmes) de' <i>us</i> minjar (femnas) minja' <i>us</i> (òmes) minja' <i>us</i> (femnas) de' <i>us</i> parlar (òmes) de' <i>us</i> parlar (femnas) parla' <i>us</i> (òmes) parla' <i>us</i> (femnas)
oc	los	PronPers3pCodApM PronPers3pCodPpM PronPers3pCoiApM (gasc) PronPers3pCoiPpM (gasc)	PronPers3pCodAp PronPers3pCodPp PronPers3pCoiAp (gasc) PronPers3pCoiPp (gasc)	<i>los</i> gaiti gaita- <i>los</i> <i>los</i> parli parla- <i>los</i>
oc	me / mi	PronPers1sCodApM PronPers1sCodApF PronPers1sCodPpM PronPers1sCodPpF	PronPers1sCodAp PronPers1sCodPp	<i>me</i> troba bèl <i>me</i> troba bèla tracta- <i>me</i> de pèc tracta- <i>me</i> de pèga

		PronPers1sCoiApM PronPers1sCoiApF PronPers1sCoiPpM PronPers1sCoiPpF PronPers1sReflApM PronPers1sReflApF PronPers1sReflPpM PronPers1sReflPpF	PronPers1sCoiAp PronPers1sCoiPp PronPers1sReflAp PronPers1sReflPp	<i>me parla</i> (òme) <i>me parla</i> (femna) <i>parla-me</i> (òme) <i>parla-me</i> (femna) <i>me pencheni</i> (òme) <i>me pencheni</i> (femna) <i>de penchenar-me</i> (òme) <i>de penchenar-me</i> (femna)
oc	m'	PronPers1sCodApM PronPers1sCodApF PronPers1sCodPpM PronPers1sCodPpF PronPers1sCoiApM PronPers1sCoiApF PronPers1sCoiPpM PronPers1sCoiPpF PronPers1sReflApM PronPers1sReflApF PronPers1sReflPpM (gasc) PronPers1sReflPpF (gasc)	PronPers1sCodAp PronPers1sCodPp PronPers1sCoiAp PronPers1sCoiPp PronPers1sReflAp PronPers1sReflPp	<i>m'apèra</i> « pèc » <i>m'apèra</i> « pèga » <i>retròba-m'i</i> (òme) <i>retròba-m'i</i> (femna) <i>m'explica</i> (òme) <i>m'explica</i> (femna) <i>explica-m'ac</i> (òme) <i>explica-m'ac</i> (femna) <i>m'èri penchenat</i> <i>m'èri penchenada</i> <i>de penchenar-m'i</i> (òme) <i>de penchenar-m'i</i> (femna)
oc	'm (gasc)	PronPers1sCodApM PronPers1sCodApF PronPers1sCodPpM PronPers1sCodPpF PronPers1sCoiApM PronPers1sCoiApF PronPers1sCoiPpM PronPers1sCoiPpF PronPers1sReflApM PronPers1sReflApF PronPers1sReflPpM PronPers1sReflPpF	PronPers1sCodAp PronPers1sCodPp PronPers1sCoiAp PronPers1sCoiPp PronPers1sReflAp PronPers1sReflPp	<i>de'm trobar</i> bèl <i>de'm trobar</i> bèla <i>tracta'm</i> de pèc <i>tracta'm</i> de pèga <i>de'm parlar</i> (òme) <i>de'm parlar</i> (femna) <i>parla'm</i> (òme) <i>parla'm</i> (femna) <i>que'm pencheni</i> (òme) <i>que'm pencheni</i> (femna) <i>de penchenà'm</i> (òme) <i>de penchenà'm</i> (femna)
oc	nos / nse	PronPers1pCodApM PronPers1pCodApF PronPers1pCodPpM PronPers1pCodPpF PronPers1pCoiApM PronPers1pCoiApF PronPers1pCoiPpM PronPers1pCoiPpF PronPers1pReflApM PronPers1pReflApF PronPers1pReflPpM PronPers1pReflPpF	PronPers1pCodAp PronPers1pCodPp PronPers1pCoiAp PronPers1pCoiPp PronPers1pReflAp PronPers1pReflPp	<i>nos tròba bèls</i> <i>nos tròba bèlas</i> <i>tracta-nos</i> de pècs <i>tracta-nos</i> de pègas <i>nos parla</i> (òmes) <i>nos parla</i> (femnas) <i>parla-nos</i> (òmes) <i>parla-nos</i> (femnas) <i>nos penchenam</i> (òmes) <i>nos penchenam</i> (femnas) <i>penchenam-nos</i> (òme) <i>penchenam-nos</i> (femnas)
oc	ns' (gasc)	PronPers1pCodApM PronPers1pCodApF PronPers1pCodPpM PronPers1pCodPpF PronPers1pCoiApM PronPers1pCoiApF	PronPers1pCodAp PronPers1pCodPp PronPers1pCoiAp	<i>ns'apèra</i> « pècs » <i>ns'apèra</i> « pègas » <i>retròba-ns'i</i> (òmes) <i>retròba-ns'i</i> (femnas) <i>ns'explica</i> (òme) <i>ns'explica</i> (femna)

		PronPers1pCoiPpM PronPers1pCoiPpF PronPers1pReflApM PronPers1pReflApF PronPers1pReflPpM PronPers1pReflPpF	PronPers1pCoiPp PronPers1pReflAp PronPers1pReflPp	explica-ns'ac (òme) explica-ns'ac (femna) ns'èrem penchenats ns'èrem penchenadas de penchenar-ns'i (òme) de penchenar-ns'i (femna)
oc	'ns (gasc)	PronPers1pCodApM PronPers1pCodApF PronPers1pCodPpM PronPers1pCodPpF PronPers1pCoiApM PronPers1pCoiApF PronPers1pCoiPpM PronPers1pCoiPpF PronPers1pReflApM PronPers1pReflApF PronPers1pReflPpM PronPers1pReflPpF	PronPers1pCodAp PronPers1pCodPp PronPers1pCoiAp PronPers1pCoiPp PronPers1pReflAp PronPers1pReflPp	de'ns trobar bès de'ns trobar bèlas tracta'ns de pècs tracta'ns de pègas de'ns parlar (òme) de'ns parlar (femna) parla'ns (òme) parla'ns (femna) que'ns penchenam (òme) que'ns penchenam (femna) de penchenà'ns (òme) de penchenà'ns (femna)
oc	n(os)aut(re)s	PronPers1pTonM	PronPers1pTon	<i>nosautres</i> , parlam ; parla de <i>nosautres</i>
oc	nosautres-meteisses...	PronPers1sReflTonM	PronPers1sReflTon	o fasèm <i>nosautres-meteisses</i>
oc	nosautras-meteissas...	PronPers1sReflTonF	PronPers1sReflTon	o fasèm <i>nosautras-meteissas</i>
oc	n(os)aut(ra)s	PronPers1pTonF	PronPers1pTon	<i>nosautras</i> , parlam ; parla de <i>nosautras</i>
oc	òm	PronPersIndefSuj	PronPersIndefSuj	òm ditz que
oc	se	PronPers3sReflApM PronPers3sReflApF PronPers3sReflPpM PronPers3sReflPpF PronPers3pReflApM PronPers3pReflApF PronPers3pReflPpM (gasc) PronPers3pReflPpF (gasc) PronPersIndefReflTon	PronPers3sReflAp PronPers3sReflPp PronPers3pReflAp PronPers3pReflPp PronPersIndefReflTon	se penchena (òme) se penchena (femna) de penchenar-se (òme) de penchenar-se (femna) <i>se</i> penchenan (òmes) <i>se</i> penchenan (femnas) de penchenar-se(òmes) de penchenar-se (femnas) pensar a se
oc	se-meteis...	PronPersIndefReflTon	PronPersIndefReflTon	o far <i>se-meteis</i>
oc	s'	PronPers3sReflApM PronPers3sReflApF PronPers3sReflPpM PronPers3sReflPpF PronPers3pReflApM PronPers3pReflApF PronPers3pReflPpM PronPers3pReflPpF	PronPers3sReflAp PronPers3sReflPp PronPers3pReflAp PronPers3pReflPp	s'èra penchenat s'èra penchenada de penchenar-s'i (òme) de penchenar-s'i (femna) s'èran penchenats s'èran penchenadas de penchenar-s'i (òmes) de penchenar-s'i (femnas)
oc	's (gasc)	PronPers3sReflApM PronPers3sReflApF	PronPers3sReflAp	de's penchenar (òme) de's penchenar (femna)

		PronPers3sReflPpM PronPers3sReflPpF PronPers3pReflApM PronPers3pReflApF PronPers3pReflPpM PronPers3pReflPpF	PronPers3sReflPp PronPers3pReflAp PronPers3pReflPp	de penchenà's (òme) de penchenà's (femna) de's penchenar (òmes) de's penchenar (femnas) de penchenà's (òmes) de penchenà's (femnas)
oc	si	PronPersIndefReflTon	PronPersIndefReflTon	òm pensa a si
oc	te / ti	PronPers2sCodApM PronPers2sCodApF PronPers2sCodPpM PronPers2sCodPpF PronPers2sCoiApM PronPers2sCoiApF PronPers2sCoiPpM PronPers2sCoiPpF PronPers2sReflApM PronPers2sReflApF PronPers2sReflPpM PronPers2sReflPpF	PronPers2sCodAp PronPers2sCodPp PronPers2sCoiAp PronPers2sCoiPp PronPers2sReflAp	te troba bèl te troba bèla tracta-te de pèc tracta-te de pèga te parla (òme) te parla (femna) parla-te (òme) parla-te (femna) te penchenas (òme) te penchenas (femna) de penchenar-te (òme) de penchenar-te (femna)
oc	t'	PronPers2sCodApM PronPers2sCodApF PronPers2sCodPpM PronPers2sCodPpF PronPers2sCoiApM PronPers2sCoiApF PronPers2sCoiPpM PronPers2sCoiPpF PronPers2sReflApM PronPers2sReflApF PronPers2sReflPpM (gasc) PronPers2sReflPpF (gasc)	PronPers2sCodAp PronPers2sCodPp PronPers2sCoiAp PronPers2sCoiPp PronPers2sReflAp	t'apèra « pèc » t'apèra « pèga » retrobatz-t'i (òme) retrobatz-t'i (femna) t'explica (òme) t'explica (femna) explicatz-t'ac (òme) explicatz-t'ac (femna) t'eras penchenat t'eras penchenada de penchenar-t'i (òme) de penchenar-t'i (femna)
oc	't (gasc)	PronPers2sCodApM PronPers2sCodApF PronPers2sCodPpM PronPers2sCodPpF PronPers2sCoiApM PronPers2sCoiApF PronPers2sCoiPpM PronPers2sCoiPpF PronPers2sReflApM PronPers2sReflApF PronPers2sReflPpM PronPers2sReflPpF	PronPers2sCodAp PronPers2sCodPp PronPers2sCoiAp PronPers2sCoiPp PronPers2sReflAp	de't trobar bèl de't trobar bèla tractatz't de pèc tractatz't de pèga de't parlar (òme) de't parlar (femna) parlà't (òme) parlà't (femna) que't penchenas (òme) que't penchenas (femna) de penchenà't (òme) de penchenà't (femna)
oc	tu	PronPers2sTonM PronPers2sTonF	PronPers2sTon	tu, parlas ; parla de tu (òme) tu, parlas ; parla de tu (femna)
oc	tu-meteis...	PronPers2sReflTonM	PronPers2sReflTon	o fas tu-meteis
oc	tu-meteissa...	PronPers2sReflTonF	PronPers2sReflTon	o fas tu-meteissa

oc	vos / vse	PronPers2pCodApM PronPers2pCodApF PronPers2pCodPpM PronPers2pCodPpF PronPers2pCoiApM PronPers2pCoiApF PronPers2pCoiPpM PronPers2pCoiPpF PronPers2pReflApM PronPers2pReflApF PronPers2pReflPpM PronPers2pReflPpF PronPers2pTonM PronPers2pTonF	PronPers2pCodAp PronPers2pCodPp PronPers2pCoiAp PronPers2pCoiPp PronPers2pReflAp PronPers2pReflPp PronPers2pTon	vos tràba bèls vos tràba bèlas tracta-vos de pècs tracta-vos de pègas vos parla (òmes) vos parla (femnas) parla-vos (òmes) parla-vos (femnas) vos penchenatz (òmes) vos penchenatz (femnas) penchenatz-vos (òme) penchenatz-vos (femnas) vos, sètz bèl ; parla de vos vos, sètz bèla ; parla de vos
oc	vos-meteis...	PronPers2pReflTonM	PronPers2sReflTon	o fasètz vos-meteis
oc	vos-meteissa...	PronPers2pReflTonF	PronPers2sReflTon	o fasètz vos-meteissa
oc	vs' (gasc)	PronPers2pCodApM PronPers2pCodApF PronPers2pCodPpM PronPers2pCodPpF PronPers2pCoiApM PronPers2pCoiApF PronPers2pCoiPpM PronPers2pCoiPpF PronPers2pReflApM PronPers2pReflApF PronPers2pReflPpM PronPers2pReflPpF	PronPers2pCodAp PronPers2pCodPp PronPers2pCoiAp PronPers2pCoiPp PronPers2pReflAp PronPers2pReflPp	vs'apèra « pècs » vs'apèra « pègas » retròba-vs'i (òmes) retròba-vs'i (femnas) vs'explica (òme) vs'explica (femna) explica-vs'ac (òme) explica-vs'ac (femna) vs'èretz penchenats vs'èretz penchenadas de penchenar-vs'i (òme) de penchenar-vs'i (femna)
oc	'vs (gasc)	PronPers2pCodApM PronPers2pCodApF PronPers2pCodPpM PronPers2pCodPpF PronPers2pCoiApM PronPers2pCoiApF PronPers2pCoiPpM PronPers2pCoiPpF PronPers2pReflApM PronPers2pReflApF PronPers2pReflPpM PronPers2pReflPpF	PronPers2pCodAp PronPers2pCodPp PronPers2pCoiAp PronPers2pCoiPp PronPers2pReflAp PronPers2pReflPp	de'vs trobar bèls de'vs trobar bèlas tracta'vs de pècs tracta'vs de pègas de'vs parlar (òme) de'vs parlar (femna) parla'vs (òme) parla'vs (femna) que'vs penchenam (òme) que'vs penchenam (femna) de penchenà'vs (òme) de penchenà'vs (femna)
oc	v(os)aut(re)s	PronPers2pTonM	PronPers2pTon	vosautres, sètz bèls ; parla de vosautres
oc	vosautres-meteisses...	PronPers2pReflTonM	PronPers2sReflTon	o fasètz vosautres-meteisses
oc	v(os)aut(ra)s	PronPers2pTonF	PronPers2pTon	vosautras, sètz bèlas ; parla de vosautras
oc	vosautras-	PronPers2pReflTonF	PronPers2sReflTon	o fasètz vosautras-meteissas

	meteissas...		
--	--------------	--	--

Tableau explicatif pour les codes de catégorie détaillée des pronoms personnels occitans :

pers.	COD PronPers**Cod	COI PronPers**Coi	COD/COI réfléchi ***0015020*2	Sujet ***0115000* 1	COI ***0315010*1	COI réfléchi ***0315020*1
Exemple	la gaita	li balha	se penchena	ela , parla	parla d' ela	pensar a se
1 sg m 111*****0*		me / mi PD111061500002		ieu PD111001500001		
1 sg f 121*****0*		me / mi PD121061500002		ieu PD121001500001		
2 sg fam m 211*****2*		te / ti PD211061500022		tu PD211001500021		
2 sg fam f 221*****2*		te / ti PD221061500022		tu PD221001500021		
2 sg poli m 211*****1*			VOS PD211001500010			
2 sg poli f 221*****1*			VOS PD221001500010			
3 sg m 311*****0*	lo PD311041501002	li (lo gasc) PD311031501002	se PD31100150200 0	el / eth PD311001501001	se (fusionné) si PD3110315020 01	
3 sg f 321*****0*	la PD321041501002	li (la gasc) PD321031501002	se PD32100150200 0	ela / era PD321001501001	se (fusionné) si PD3210315020 01	
1 pl m 112*****0*		nos PD11261501002		n(os)aut(re)s PD112001501001		
1 pl f 122*****0*		nos PD122061501002		n(os)aut(r)as PD122001501001		
2 pl m 212*****0*		vos PD212061501002		v(os)aut(re)s PD212001501001		
2 pl f 222*****0*		vos PD222061501002		v(os)aut(r)as PD222001501001		
3 pl m 312*****0*	los / lei(s) PD312041501002	lor / li (los gasc) PD312031501002	se PD31200150200 0	eles / eths PD312001501001	se (fusionné) si PD3120315020 01	
3 pl f 322*****0*	las / lei(s) PD32204150100 2	lor / li (las gasc) PD32203150100 2	se PD3220015020 00	elas / eras PD322001501001	se (fusionné) si PD3220315020 01	

Tableau explicatif pour les codes de catégorie détaillée des pronoms personnels français :

	Pronoms conjoints/atones					Pronoms disjoints/toniques				
pers.	Sujet ***011500 0*2	COD anteposé ***0415 010*2	COI anteposé ***03150 10*2	COD/COI réfléchi ***001502 0*2	COD/COI postposé ***0415000 *2 ***0315000 *2	Sujet ***011500 0*1	COI ***0315 010*1	COI réfléchi ***0315020 1		
Ex.	elle parle	il la regarde	il lui donne	elle se peigne	dis-lui	elle, elle parle	il parle d'elle	penser à soi		
1 sg m 111*****0*	je PD3120115 00002	me PD111041 500002	me PD111031 500002		moi PD111041500002 PD111031500002	moi PD111001500001				
1 sg f 121*****0*	je PD1210115 00002	me PD12104 150000 2	me PD121031 500002		moi PD121041500002 PD121031500002	moi PD121001500001				
2 sg fam m 211*****2*	tu PD2110115 00022	te PD21104 1500022	te PD211031 500022		toi PD211041500022 PD211031500022	toi PD211001500021				
2 sg fam f 221*****2*	tu PD2210115 00022	te PD22104 1500022	te PD22103 1500022		toi PD221041500022 PD221031500022	toi PD221001500021				
2 sg poli m 211*****1*					vous PD211001500010					
2 sg poli f 221*****1*					vous PD221001500010					
3 sg m 311*****0*	il PD3110115 00002	le PD31104 1501002	lui PD311031 501002	se PD3110015 02002		lui PD311001501001	soi PD31103150 2001			
3 sg f 321*****0*	elle PD3210115 00002	la PD32104 1501002	lui PD32103 1501002	se PD3210015 02002		elle PD321001501001	soi PD32103150 2001			
1 pl m 112*****0*					nous PD112001500000					
1 pl f 122*****0*					nous PD122001500000					
2 pl m 212*****0*					vous PD212001500000					
2 pl f 222*****0*					vous PD222001500000					
3 pl m 312*****0*	ils PD3120115 00002	les PD31204 1501002	leur PD31203 1501002	se PD3120015 02002		eux PD312001501001	soi PD31203150 2001			
3 pl f 322*****0*	elles PD322011 500002	les PD32204 4150100 2	leur PD32203 1501002	se PD322001 502002		elles PD322001501001	soi PD3220315 02001			

7 Annexes

7.1 Codes langue de l'IANA (balises `<usg type="lang">`, attributs `xml:lang`)

Les codes langue et région utilisés sont ceux du registre de l'IANA :

<http://www.iana.org/assignments/language-subtag-registry/language-subtag-registry>

Pour l'occitan, on utilise le code « oc », complété par un subtag de variété s'il ne s'agit pas d'occitan commun, et par un subtag de graphie. Attention, on ne peut mettre qu'un subtag de chaque catégorie.

Subtag de variétés (à ajouter après `oc` suivi d'un -) :

Variété	Subtag
Auvergnat	auvern
Gascon	gascon
Limousin	lemosin
Languedocien	lengadoc
Provençal	provenc
Vivaro-alpin	vivaraup
Variétés du Croissant	creiss
Aranaïs	aranes
Cisalpin	cisaup
Nissart	nicard

Subtag de graphie (à faire précéder d'un -) : suit la balise `oc` éventuellement complétée d'un subtag de graphie :

Graphie	Subtag
Graphie classique	grclass
Graphies mistralienne et dérivées (fébusienne, <i>escolo dou Po...</i>)	grmistr
Graphie italianisante	grital
Autres graphies	grautes (subtag non déclaré à l'IANA)

Exemples :

- Occitan gascon en graphie fébusienne : `oc-gascon-grmistr`
- Occitan languedocien en graphie par défaut/non précisée : `oc-lengadoc`
- Occitan commun (graphie classique) : `oc-grclass`

- Occitan, variété non précisée, graphie classique : **oc-grclass**

Dans les données du Congrès, on ne trouve que des codes sous la forme « fr », « oc-graphie » ou « oc-variété-graphie ».

7.2 Comment choisir entre unité grammaticale et phrase

Les **<form>** contenant plusieurs mots peuvent être étiquetées en tant qu'unité grammaticale (**word**) ou en tant que **phrase**.

Une unité grammaticale peut être un mot seul, mais aussi d'un groupe de mots non réductibles, grammaticalement, à la somme de ses parties. Par exemple, l'adverbe « fa pas gaire » (ex. : « i anèri fa pas gaire ») se conduit de manière différente, grammaticalement, de far (verbe) + pas (adv. neg) + gaire (adv.) (ex. : « pauc de blat fa pas gaire de farina »).

Pour savoir lequel choisir, on regarde la catégorie de la forme dans son ensemble et on procède comme décrit ci-dessous.

On étiquette comme unité grammaticale les séries de mot dont le tout n'est pas égal, grammaticalement, à la somme des parties. Si, grammaticalement, la série de mot se comporte pareil qu'elle soit une unité grammaticale ou une somme de mots, alors on l'étiquette **phrase**. Si la série de mots a un fonctionnement particulier, alors il faut l'étiqueter unité grammaticale.

Pour savoir si le tout est égal à la somme des parties, on peut :

- Voir si d'autres mots au hasard avec la même catégorie fonctionneraient pareil
- Voir si la fonction d'une traduction (composée de mots d'autres catégories) serait la même

Exemple :

Ex. 1: comparaison de traductions	pèira levada n. + adj. = groupe nominal	dolmen n. = groupe nominal	ok
Ex. 1: autres mots au hasard	pèira levada n. + adj.	cat polit n. + adj. La tournure est possible et est utilisée dans les mêmes contextes.	ok
Ex. 2 : comparaison de traductions	a plaser prép + n. = compl. du nom	doucement adv. = groupe adverbial	pas bon
Ex. 2 : autres mots au hasard	a plaser prép + n.	de veitura prép + n. Ce genre de tournure n'est pas du tout utilisée dans le même contexte.	pas bon

7.2.1 De manière générale

On étiquette comme **unité grammaticale**:

- Si certains des mots qui composent la forme n'existent pas en tant qu'unité (ex. : « priori » dans « a priori », « mobil » et « home » dans « mobil home »...).
- Si l'unité grammaticale existe avec une autre catégorie grammaticale (ex. : « cap e cap » nom commun qui existe aussi comme « cap e cap » adjectif)
- S'il s'agit d'une conjonction, d'un déterminant, d'un numéral, d'une préposition ou d'un pronom.
- S'il s'agit d'une interjection qui ne contient pas de verbe conjugué (ex. : « Bonsoir ! »)
- S'il s'agit d'une forme en un seul mot avec une ou plusieurs collocations (ex. : « se vêtir de »).
- Les noms communs composés d'un verbe et un nom (ex. : « garda nèns ») ou d'un verbe et d'un adjectif (ex. : « canta polit ») ou d'un adverbe et d'un nom (ex. : « hors jeu »)

On étiquette comme **phrase**:

- Les verbes et les noms (hormis ceux listés ci-dessus).
- Les interjections qui contiennent un verbe conjugué (ex. : « Dieu te bénisse ! ») ou de type quel(le) + nom (ex. : « quel culot ! ») et possessif + nom (ex. : « mon coeur », « votre dévoué »...) ou qui contiennent une ponctuation (ex. : « tè, tu ! »)

7.2.2 Adverbes

On étiquette comme phrase les adverbes composés comme suit :

- préadverbe + adverbe (ex. : « tout fort », « très bas », « quasi rien »...) sauf si l'adverbe peut aussi être un modal ou un conjonctif (ex. : « tout d'abord »)
- adverbe + et/ou + adverbe (ex. : « à tort et à travers », « ici ou ailleurs »...)
- selon + nom (ex. : « selon moi »)
- COI lexicalisés (ex. : « au plus offrant », « au premier venu »...)
- compléments circonstanciels de lieu composés comme suit :
 - préposition + déterminant + nom (ex. : « sur le devant », « au milieu », « à la maison », « par les chemins »...) sauf « au ras », « en l'aire », « à l'intérieur », « à l'extérieur ».
 - préposition + pronom (ex. : « devant moi », « avec personne »...)
 - en/dans + lieu
 - en + déterminant + nom (ex. : « en plusieurs endroits »)
 - au gré de + nom (ex. : « au gré de l'eau »)
 - introduits par chez (ex. : « chez lui »)
 - finissant par estant (ex. : « de casa estant »)
 - près de + lieu (ex. : « près d'ici », « près du lac »)
 - chez (+ déterminant) + nom/pronom (ex. : « chez moi », « chez le coiffeur »...)
 - par + moment (ex. : « par le passé », « par un beau dimanche »...)
 - vers + lieu/direction (ex. : « vers la maison », « vers l'ouest »)
 - là où + proposition (ex. : « là où je vais », « là où elles vivent »...)
 - au début de/au milieu de/au bout de + lieu (ex. : « au milieu du chemin », « au bout du pont »...)
- compléments circonstanciels de temps composés comme suit :

- dans/au milieu de/au cœur de/pour/au moment de/au début de/à la fin de/au bout de... + nom exprimant une période (ex. : « au cœur de l'été », « pour Noël », « dans la matinée », « au moment de partir »...)
 - introduits par « au moment » (ex. : « au moment voulu », « au moment où il part »...)
 - avant de/après + proposition/infinitif (ex. : « avant de partir », « après avoir réfléchi »...); avant/après + déterminant + nom (ex. : « avant la nuit », « après l'heure »)
 - avant/après + moment (ex. : « avant 15h », « avant l'été », « après décembre »...)
 - avant de + infinitif (ex. : « avant de partir »...)
 - introduits par au temps de/du temps de/au temps que (ex. « au temps des pharaons », « au temps que Bèrta fialava »)
 - à + heure (+ tapantes/pile) (ex. : « à neuf heures », « à midi pile »...)
 - le + date (ex. : « le premier janvier »)
 - commençant par une action (ex. : « au coucher du soleil », « à la tombée de la nuit », « à l'approche du jour », « à l'entrée de l'hiver »...)
 - quand + proposition (ex. : « quand les poules auront des dents »)
 - du temps de/au(x) temps de + divers (ex. : « du temps des pharaons »)
 - autour de/pendant/dans l'affaire de/au cours de/dans le courant de/dans/après/avant + période (ex. : « autour de Noël », « pendant les vacances », « dans le courant du mois », « dans l'affaire d'une minute »...)
 - à + déterminant + numéral + nom (ex. : « à la première heure », « au troisième top »...)
 - introduits par il y a (et « fa » en occitan) (ex. : « il y a longtemps »)
 - introduits par (au)tanlèu com (ex. : « autanlèu com vienut »)
 - introduits par tous les/chaque (ex. : « tous les mois », « chaque jour »)
 - introduits par « jusqu'à »/« jusque » (ex. : « jusqu'à lundi », « jusque demain », « jusqu'à nouvel ordre »...)
 - en + durée (ex. : « en une heure », « en un clin d'oeil »)
 - vers + moment (ex. : « vers une heure », « vers Noël », « vers décembre »...)
 - pour + moment (ex. : « pour le 11 novembre », « pour lundi », « pour Noël », « pour 16h »)
 - sur le coup de/autour de + h (ex. : « autour de 4h », « sur le coup de midi »...)
 - depuis que + proposition (ex. : « depuis que le monde est monde », « depuis qu'il est parti »...)
 - quand + proposition (ex. : « quand j'arriverai », « quand les poules auront des dents »...)
 - à partir de + moment (ex. : « à partir de maintenant », « à partir de jeudi »...)
- compléments circonstanciels de manière composés comme suit :
- à la/à la mode/manière (de) + nom propre ou nationalité (ex. : « à la mode turque », « à l'anglaise », « à la Marengo »...)
 - pour la + ordinal + fois (« pour la troisième fois »...)
 - introduits par « au fil de » (ex. : « au fil de l'eau »)
 - introduits par comme (ex. : « comme dans un rêve », « comme ci comme ça », « comme si de rien n'était »...)
 - en + participe présent (ex. : « en chantant »).
 - en (= en tant que) + nom (ex. : « en connaisseur »)

- en + langue
- pour + proposition/infinitif (ex. : « pour rire », « pour qui en veut »...)
- sans + verbe (ex. : « sans bouger »)
- avec/sans + nom (ex. : « avec animation », « sans douceur »)
- introduits par « à l'état de » (ex. : « laisser à l'état de brouillon »)
- déterminant + partie du corps + adjectif (ex. : « la tête nue », « les bras ballants »...)
- adverbes introduits par à titre/au nom de (ex. : « à titre privé », « à titre posthume », « à mon nom », « au nom de la loi »)
- a + tot + infinitif (ex. : « a tot pèrdré », « a tot petar »...)
- aux dépens de/aux frais de (ex. : « aux frais de la princesse »)
- au bon vouloir de/au gré de (ex. : « à mon gré », « au bon vouloir de Monsieur »...)
- introduits par comme (ex. : « comme un dératé », « comme s'il en pleuvait »)
- introduits par au train où (ex. : « au train où vont les choses »)
- verbe 3e pers. + que + verbe 3e pers. (ex. : « coûte que coûte », « vaille que vaille »...)
- par + déterminant + position/passage (ex. : « par le haut », « par le dessus », « par la fenêtre »...)
- par + moyen (ex. : « par avion »)
- par + article + partie du corps (ex. : « par la main »)
- au son de + instrument
- aux + proportion (ex. : « aux trois quarts »)
- introduits par « d'un air »
- compléments circonstanciels de but
- si + proposition (ex. : « si j'étais toi »)
- selon + personne/prénom personnel (ex. : « selon Confucius », « selon moi »...)
- introduits par « en vertu de »
- proportion construite avec « par » + nom (ex. : « par habitant », « par an »...) sauf « pour cent » et « pour mille ».
- prix, distances, poids (ex. : « à 15 euros », « dans 10km », « pour 10 kilos »)
- à + verbe (ex. : à crever)
- en + matière/langue (ex. : « en bois », « en anglais »)
- en + mesure/monnaie (ex. : « en mètres », « en euros »)
- en + état/situation/aspect (ex. : « en voyage », « en retard », « en colère », « en bleu », « en italique », « en poudre », « en activité »...)
- en qui introduit un COI (ex. : « croire en quelqu'un »)
- en + discipline/domaine (ex. : « expert en médecine », « un requin en affaires », « fidèle en amour »)
- de + nom + en + nom (ex. : « de ville en ville », « de fil en aiguille »)
- hors + substantif (ex. : « hors cadre », « hors contexte »...)
- par + cause (ex. : « par inattention », « par hasard ») hormis les adverbes modaux
- distribution introduite par « par » (ex. : « par semaine », « par personne »...)
- par + pronom personnel (ex. : « par moi-même »)
- malgré + pronom / malgré + art + nom (ex. : « malgré moi », « malgré les circonstances »...)
- pour + personne (dans le sens de « de l'avis de » + personne)
- si + verbe conjugué (« si cela arrive », « s'il faut »...)

- selon/d'après + proposition (« d'après ce qu'on dit », « selon s'il pleut »...)

Les autres adverbes sont étiquetés comme unité grammaticale.

7.2.3 Adjectifs

On étiquette comme unité grammaticale les adjectifs composés comme suit :

- à + article défini + nom ne correspondant pas à un lieu (ex. : « à l'arrachée »)
- de + nom (ex. : « de travers »)
- en + position (ex. : « en avant »)
- adjectif de couleur + nom (ex. : « bleu roi »)
- autres adjectifs à l'exception de ceux listés ci-dessous

On étiquette comme phrase les adjectifs composés comme suit :

- à + nom dans le sens de « avec » (ex. : « à rames »)
- à + verbe (ex. : « à boire »)
- à + article défini + nom correspondant à un lieu (ex. : « à l'abri »)
- à/pour + verbe (ex. : « à vendre »)
- comme + proposition (ex. : « comme il convient »)
- déterminant éventuel + nom + adjectif (ex. : « tête nue », « les yeux marqués »...)
- dans/sur + article + nom (ex. : « dans la gêne »)
- de + bon/mauvais/excellent/meilleur/premier/même... + nom (ex. : « de bonne qualité »)
- de + bon/mal + verbe (ex. : « de bon far »)
- de + verbe : (ex. : « de plâner »)
- de + déterminant + nom (ex. : « de l'endroit », « des environs »...)
- de + matière (ex. : « de chair »)
- de + origine (ex. : « de derrière les fagots », « d'ici »...)
- de + période (ex. : « de Noël »)
- en/dans + lieu
- en + déterminant + nom (ex. : « en plusieurs endroits »)
- en + matière/langue (ex. : « en bois », « en anglais »)
- en + mesure/monnaie (ex. : « en mètres », « en euros »)
- en + état/situation (ex. : « en voyage », « en retard », « en colère », « en bleu », « en tialique », « en activité »...)
- en + moment/durée (ex. : « en hiver », « en une heure »)
- en (= en tant que) + nom (ex. : « en connisseur »)
- en + vêtement (ex. : « en chemise »)
- hors (+ de) + nom/pronom (ex. : « hors de portée », « hors de lui »...)
- sans + nom (ex. : « sans mesure »)
- qui + verbe (ex. : « qui arrive »)
- adj + groupe complément (ex. : « malin comme un singe », « libre dans ses propos », « frais arrivé »...)
- composé de plusieurs adjectifs reliés par une conjonction (ex. : « franc et quitte »)
- mal/bien + adjectif (ex. : « bien élevé »)
- adv intensité (à demi/mi/tout/presque/à moitié/peu/plus/gravement) + adjectif (ex. : « à demi-cuit », « mi-cuit », « tout petit », « presque prêt », « à moitié nu », « peu nombreux »...)
- adj de couleur + adj (ex. : « bleu pâle »)
- « couleur » + nom (ex. : « couleur crème »)

- pur + nom (ex. : « pur porc »)
- non + adjetif (ex. : « non exprimé »)
- plein de + nom (ex. : « plein de morgue »)
- nom + eishenge/eishenja (moneda eishenge)
- sur + article + nom
- comme + article + nom
- commençant par « avec »
- au temps de ...
- hors + substantif (ex. : « hors cadre », « hors contexte »...)
- que + proposition (ex. : « qu'il faut », « qu'on peut »...)

7.2.4 Noms propres

On étiquette comme unité grammaticale :

- Les noms propres composés de nom géographique + de + nom propre (ex. : « seuil de Naurouze », « chutes du Niagara »...) ou sous forme de périphrase (ex. : « la Ville éternelle », « le Nouveau Monde »...), mais ils ont la catégorie grammaticale du nom seul et le nom géographique + de est mis en collocation (ex. : <colloc>chutes du</colloc><orth>Niagara</orth><gramGrp><gram type="catDet" norm="NomPropMS"/>)

On étiquette comme phrase :

- Les noms propres qui sont des noms d'oeuvres (ex. : « Les Misérables »), sauf si cet ouvrage est si connu qu'il est passé en nom commun (ex. : « Coran », « Ancien Testament », « Code de la route »...).

Les autres noms propres sont étiquetés comme unités grammaticales.

7.3 Comment étiqueter le type des sous-entrées <re> des dictionnaires ?

7.3.1 Attestation (attestation)

Il s'agit de citations dans des œuvres historiques, ayant pour but d'attester de l'usage du mot-vedette à une époque donnée.

7.3.2 Citacion (citation)

Les citations sont extraites d'œuvres. En général, elles sont accompagnées d'un titre ou d'un nom d'auteur.

7.3.3 Collocation (collocation)

Lorsque la sous-entrée est le mot-vedette auquel on a adjoint une ou plusieurs collocations, sans autre ajout. Ex. : venir - venir de.

Attention, lorsque la collocation est un pronom et le mot vedette un verbe, la sous-entrée est en général un verbe pronominal, pas une collocation (voir 7.3.14).

Dans les cas où l'on a plusieurs collocations, dont des pronoms et des non-pronoms, on indique « collocation ». Ex. : souvenir - se souvenir de.

7.3.4 Dérivé (derived)

Les dérivés sont des mots formés à partir de la racine du mot-vedette. Ex : perfièit - perfectible.

7.3.5 Exemple (example)

Les exemples sont des suites de mots qui ne sont pas « figées ». Un mot peut être échangé avec un synonyme ou un équivalent (« se copar la camba » - « se copar lo pè »), on peut y ajouter des adverbes ou des adjectifs (« se copar la camba » - « se copar bravament la camba »). L'exemple est créé par les lexicographes qui réutilisent le mot-vedette pour montrer au lecteur comment peut être utilisé le mot expliqué.

Quand une expression est conjuguée, en général, elle est classée dans les exemples. « Córrer a hum de calhau » sera une expression tandis que « Que correishi a hum de calhau » sera un exemple, car c'est une façon d'employer l'expression.

De plus, pour un exemple, si nous donnons à quelqu'un la définition de tous les mots qui la composent, il en comprendra le sens. Par exemple, « secouer un vêtement » se comprend avec les définitions de « secouer » et de « simplement ». Par contre, « secouer les puces » ne se comprend pas si on se base uniquement sur la définition de « secouer » et de « puce ». Même chose pour « haute mer », « coup fourré »...

C'est la différence entre un exemple et une expression ou une locution. Ainsi, « secouer un vêtement » ne renvoie pas à une réalité désignée par la forme mais est une illustration de l'emploi du verbe « secouer ». A l'inverse, « secouer les puces » renvoie à une image, une idée.

7.3.6 Expression (expression)

Contrairement à l'exemple, une expression est figée (avec différents degrés de figement, elle peut ne pas l'être complètement). On ne peut généralement pas y échanger un mot avec un de ses synonymes ou équivalents (« arcuéller com un can en ua glèisa » - « *arcuéller com un gat en ua glèisa ») ou y ajouter un adverbe (« arcuéller com un can en ua glèisa » - « *arcuéller com un gròs can en ua glèisa »).

Une expression est imagée. Souvent, on ne peut pas comprendre une expression même si on a une définition de tous les mots qui la composent. Par exemple, « secouer les puces » ne se comprend pas si on se base uniquement sur la définition de « secouer » et de « puce ». Même chose pour « haute mer », « coup fourré »...

Quand une expression est conjuguée, en général, elle est classée dans les exemples. « Córrer a hum de calhau » sera une expression tandis que « Que correishi a hum de calhau » sera un exemple, car c'est une façon d'employer l'expression.

En résumé, une expression est comme une locution, mais avec la notion d'image (métaphore, comparaison).

7.3.7 Gentilé (demonym)

Lorsque la sous-entrée est le nom des habitants du mot-vedette (qui est un lieu). Ex. : Toulouse - toulousains.

7.3.8 Interjection (interjection)

Une interjection est généralement accompagnée d'un point d'exclamation.

Attention, un ordre n'est pas une interjection. Les interjections ne sont pas littérales. Si je dis « Tais-toi ! », je veux que la personne se taise. C'est un exemple d'utilisation du verbe « taire ». Par contre, si je dis « Diu t'assiste ! » à quelqu'un qui éternue, je ne veux pas forcément que Dieu l'assiste. C'est une interjection figée et autonome.

Attention, si l'interjection est une onomatopée, elle est étiquetée comme onomatopée et pas comme interjection (voir 7.3.12).

7.3.9 Locution (locution)

La différence entre une locution et un exemple est la même qu'entre une expression et un exemple (voir en 7.3.6).

La différence entre une locution et une expression est que l'expression fait appel à une image (métaphore, comparaison), alors qu'une locution est davantage une tournure avec un sens non figuré. Ex. : « avoir faim » est une locution (sens propre) alors que « avoir une faim de loup » est une expression (elle fait appel à une image).

Les mots composés du mot-vedette et d'autres mots dont il est séparé par un espace (ex. : « ajonc » -> « ajonc nain ») sont considérés comme des locutions nominales et sont classés dans la catégorie « locutions ».

7.3.10 Mot composé (compound)

Les mots composés sont des mots formés à partir du mot-vedette, en lui adjoignant un ou plusieurs autres mots auxquels il est collé (pòrtaparaula) ou dont il est séparé par un tiret (pòrta-paraula).

Lorsque le mot-vedette est séparé des autres mots par un espace (vessina de lop), on considère qu'il s'agit d'une locution nominale et on classe la forme dans les locutions.

7.3.11 Nom propre (proper)

Les noms de communes, les noms liés à la religion, les constellations... sont classés dans la catégorie « noms propres ».

7.3.12 Onomatopée (onomatopeia)

Lorsque l'expression est une onomatopée. Ex. : canard - coin-coin.

7.3.13 Proverbe (proverb)

Ce sont des sentences, il y a une sorte de sagesse, de morale à l'intérieur. Ils apportent de la « connaissance ».

7.3.14 Verbe pronominal (pronominal)

Tous les verbes en « s' », « se »... Ex : « alunhar » - « s'alunhar »..

8 Modèle de document

```
<TEI>
  <teiHeader>
    <fileDesc>
      <titleStmt>
        <title level="m" type="main">Titre de l'œuvre électronique</title>
        <title level="m" type="subc">Sous-titres éventuels</title>*
        <title level="m" type="subc">Format TEI</title>
        <title level="m" type="short">Nom du document dans l'interface du site du Congrès</title>
        <author nymRef="Nom de famille pour l'ordre" (key="ISNI XXXX") xml:lang="fr">Prénom Nom de
l'auteur en français</author>*
        <author nymRef="Nom de famille pour l'ordre" (key="ISNI XXXX") xml:lang="oc">Prénom Nom de
l'auteur en occitan</author>*
        <respStmt>
          <resp xml:lang="fr">Action en français</resp>
          <resp xml:lang="oc">Action en occitan</resp>
          <name nymRef="Nom de famille pour l'ordre" (key="ISNI XXXX") xml:lang="fr">Prénom Nom
(éventuellement organisme)</name>
          <name nymRef="Nom de famille pour l'ordre" (key="ISNI XXXX") xml:lang="oc">Prénom Nom
(éventuellement organisme)</name>
        </respStmt>*
      </titleStmt>
      <editionStmt>
        <edition>Numéro de l'édition électronique</edition>
      </editionStmt>*
      <publicationStmt>
        <authority xml:lang="fr">Organisme responsable de la diffusion (en français)</authority>
        <authority xml:lang="oc">Organisme responsable de la diffusion (en occitan)</authority>
        <availability status="free/restricted/unknown">
          <p xml:lang="fr">Droits explicites en français</p>
          <p xml:lang="oc">Droits explicites en occitan</p>
          <licence target="lien"></licence>
        </availability>
        <idno type="documentID">Pseudo unique du document</idno>
        <date when="année sur 4 chiffres" xml:lang="fr">Date explicite en français</date>
        <date when="année sur 4 chiffres" xml:lang="oc">Date explicite en occitan</date>
      </publicationStmt>
      <seriesStmt>
        <title type="main" level="s" subtype="project" xml:lang="fr">Ressources TEI du Congrès</title>
        <title type="main" level="s" subtype="project" xml:lang="oc">Ressorsas TEI deu Congrès</title>
        <title type="main" level="s" subtype="division" xml:lang="fr"
key="forms/flexions/validation/coordinates/pronunciations/lexical_bases/translations/expressions/
monolingual_lexicons/mononlingual_corpus/bilingual_corpus/historical_lexicons]">Nom en français de la
division : Lexiques de formes / Lexiques de formes fléchies / Lexiques de validation / Lexiques de
coordonnées / Lexiques de prononciations / Basas de donadas lexicalas / Dictionnaires de traductions /
Dictionnaires d'expressions / Dictionnaires monolingues / Corpus monolingues / Corpus parallèles /
Dictionnaires historiques...</title>
        <title type="main" level="s" subtype="division" xml:lang="oc"
key="forms/flexions/validation/coordinates/pronunciations/lexical_bases/translations/expressions/
monolingual_lexicons/mononlingual_corpus/bilingual_corpus/historical_lexicons]">Nom en occitan de la
division : Lexics de formas / Lexics de formas flechidas / Lexics de validacion / Lexics de coordonadas /
Lexics de prononciacions / Lexics de variantas e sinonimes / Diccionaris de traductions / Diccionaris
d'expressions / Diccionaris monolingües / Còrpus monolingües / Còrpus parallèles / Diccionaris
istorics...</title>
        <title type="main" level="s" subtype="collection" xml:lang="fr">Nom en français de la collection :
tèrm'Òc / Literatura Gascona / Revue Reclams...</title>+
        <title type="main" level="s" subtype="collection" xml:lang="oc">Nom en occitan de la collection :
tèrm'Òc / Literatura Gascona / Revista Reclams...</title>+
      </seriesStmt>
      <notesStmt>
```

```

<note rend="visible" xml:lang="fr">Notes en français sur l'aspect de la ressource électronique</note>
<note rend="visible" xml:lang="oc">Notes en occitan sur l'aspect de la ressource électronique</note>
</notesStmt>?
<sourceDesc>
  <biblStruct>
    <analytic>
      <title type="main" level="a">Article, poème, lexique extrait d'un livre ou journal décrit dans
      "monogr"</title>
      <title level="m" type="subc">Sous-titres éventuels</title>*
        <author nymRef="Nom de famille pour l'ordre" (key="ISNI XXXX") xml:lang="fr">Nom de l'auteur
        en français</author>*
          <author nymRef="Nom de famille pour l'ordre" (key="ISNI XXXX") xml:lang="oc">Nom de l'auteur
        en occitan</author>*
          <respStmt>
            <resp xml:lang="fr">Action en français</resp>
            <resp xml:lang="oc">Action en occitan</resp>
            <name nymRef="Nom de famille pour l'ordre" xml:lang="fr" (key="ISNI XXXX")>Prénom Nom
            (éventuellement organisme)</name>
            <name nymRef="Nom de famille pour l'ordre" xml:lang="oc" (key="ISNI XXXX")>Prénom Nom
            (éventuellement organisme)</name>
          </respStmt>*
        <textLang mainLang="oc-lengadoc" otherLangs="fr" xml:lang="fr">Description de la / des
        langue(s) du texte en français</textLang>
        <textLang mainLang="oc-lengadoc" otherLangs="fr" xml:lang="oc">Description de la / des
        langue(s) du texte en occitan</textLang>
        <note rend="visible" xml:lang="fr">Autre information en français</note>?
        <note rend="visible" xml:lang="oc">Autre information en occitan</note>?
    </analytic>?
    <monogr>
      <title level="m" type="main">Titre complet de l'imprimé, forme normalisée</title>
      <title level="m" type="subc">Sous-titres éventuels</title>*
      <biblScope unit="volume">Numéro du volume<biblScope>?
      <biblScope unit="issue">Numéro de revue<biblScope>?
      <author nymRef="Nom de famille pour l'ordre" (key="ISNI XXXX") xml:lang="fr">Nom de l'auteur
      en français</author>*
        <author nymRef="Nom de famille pour l'ordre" (key="ISNI XXXX") xml:lang="oc">Nom de l'auteur
      en occitan</author>*
        <respStmt>
          <resp xml:lang="fr">Action en français</resp>
          <resp xml:lang="oc">Action en occitan</resp>
          <name nymRef="Nom de famille pour l'ordre" (key="ISNI XXXX") xml:lang="fr">Prénom Nom
          (éventuellement organisme)</name>
          <name nymRef="Nom de famille pour l'ordre" (key="ISNI XXXX") xml:lang="oc">Prénom Nom
          (éventuellement organisme)</name>
        </respStmt>*
      <textLang mainLang="oc-lengadoc" otherLangs="fr" xml:lang="fr">Description de la / des
      langue(s) du texte en français (langue principale = langue des mots-vérités pour les
      dictionnaires)</textLang>
      <textLang mainLang="oc-lengadoc" otherLangs="fr" xml:lang="oc">Description de la / des
      langue(s) du texte en occitan</textLang>
      <edition xml:lang="fr">Edition de l'ouvrage (2e, de 1988...) en français</edition>?
      <edition xml:lang="oc">Edition de l'ouvrage (2e, de 1988...) en occitan</edition>?
      <idno type="ISBN">ISBN</idno>?
      <imprint>
        <publisher xml:lang="fr">Nom de l'éditeur en français</publisher>+
        <publisher xml:lang="oc">Nom de l'éditeur en occitan</publisher>+
        <pubPlace xml:lang="fr">Lieu de publication en français</pubPlace>
        <pubPlace xml:lang="oc">Lieu de publication en occitan</pubPlace>
        <date when="année sur 4 chiffres" xml:lang="fr">Date de publication en français</date>
        <date when="année sur 4 chiffres" xml:lang="oc">Date de publication en occitan</date>
    
```

```

</imprint>
<ptr type="img" target="url de la couverture"/>?
<ptr type="url" subtype="achat" target="url pour acheter l'ouvrage"/>?
<note rend="visible" xml:lang="fr">Autre information en français</note>?
<note rend="visible" xml:lang="oc">Autre information en occitan</note>?
</monogr>
<series>
  <title level="m" type="main">Si l'ouvrage appartient à une série, collection, etc.</title>
</series>?
</biblStruct>
</sourceDesc>
</fileDesc>
<encodingDesc>
  <projectDesc>
    <p xml:lang="fr">Description en français du projet dans lequel s'inscrit l'ouvrage numérique.</p>+
    <p xml:lang="oc">Description en occitan du projet dans lequel s'inscrit l'ouvrage numérique.</p>+
  </projectDesc>?
  <editorialDecl>
    <normalization>
      <p xml:lang="fr">Principes et choix éditoriaux par rapport au document d'origine en français</p>+
      <p xml:lang="oc">Principes et choix éditoriaux par rapport au document d'origine en occitan</p>+
    </normalization>?
    <correction>
      <p xml:lang="fr">Indications de modifications et/ou corrections éventuelles faites par rapport au
      texte original, en français.</p>+
      <p xml:lang="oc">Indications de modifications et/ou corrections éventuelles faites par rapport au
      texte original, en occitan.</p>+
    </correction>?
    <interpretation>
      <p xml:lang="fr">Description de l'étiquetage morpho-syntaxique pour les corpus, en français</p>+
      <p xml:lang="oc">Description de l'étiquetage morpho-syntaxique pour les corpus, en occitan</p>+
    </interpretation>?
  </editorialDecl>
  <schemaSpec start="TEI_Lite">
    <desc xml:lang="fr">Description en français du schéma ou de la DTD utilisée</desc>
    <desc xml:lang="oc">Description en occitan du schéma ou de la DTD utilisée</desc>
  </schemaSpec>
</encodingDesc>
<profileDesc>
  <langUsage>
    <language ident="code langue(-variété-graphie)" xml:lang="fr" usage="headwords">langue (et
    variété et graphie) des mots-vedettes pour les lexiques (en français)</language>?
    <language ident="code langue(-variété-graphie)" xml:lang="oc" usage="headwords">langue (et
    variété et graphie) des mots-vedettes pour les lexiques (en occitan)</language>?
    <language ident="code langue(-variété-graphie)" xml:lang="fr" usage="translations">langue (et
    variété et graphie) des traductions pour les lexiques (en français)</language>*
    <language ident="code langue(-variété-graphie)" xml:lang="oc" usage="translations">langue (et
    variété et graphie) des traductions pour les lexiques (en occitan)</language>*
    <language ident="code langue(-variété-graphie)" xml:lang="fr">langue (et variété et graphie) des
    phrases d'un corpus (en français)</language>*
    <language ident="code langue(-variété-graphie)" xml:lang="oc">langue (et variété et graphie) des
    phrases d'un corpus (en français)</language>*
  </langUsage>
</profileDesc>
</teiHeader>
<text>
  <body>
    <entry xml:id="entry-[identifiant]" cert="["current/delete/not_ok/ok/confirmed"]">
      <ptr type="["audio/img/url/video/other]" subtype="["catégorie]" target="["url]" (targetLang="["code
      langue(-variété-graphie)]")></ptr>

```

```

<form subtype="[phrase/morpheme/symbol/word]" type="[inflected/elided/inflectelided/lemma]" xml-
lang="[code langue(-variété-graphie)]" xml:id="form-[identifiant]" (n="[identifiant homographe]")
cert="[current/not_ok/not_ok_var/ok]" value="[collocations + orthographe traitées pour la recherche]">
    <colloc type="[pre/post]" norm="[forme normalisée]" (orig="[forme originale]") (expand="[forme
pour l'édition numérique]" (subtype="[catégorie détaillée]" value="[collocation traitée pour la
recherche]">collocation usuelle</colloc>*
        <orth norm="[orthographe normalisée]" (orig="[orthographe originale]" (expand=[orthographe
pour l'édition numérique]" value="[orthographe pour la recherche]">[orthographe de la traduction]</orth>
        <pron norm="[forme normalisée]" (orig="[forme originale]" (expand=[forme pour l'édition
numérique]" extent="[part/full/pref/suff/poorRhyme/richRhyme]" (notation="[type de
notation]">pronunciation</pron>*
            <syll norm="[forme normalisée]" (orig="[forme originale]" (expand=[forme pour l'édition
numérique]" extent="[part/full/pref/suff]">découpage en syllabes</syll>*
                <form subtype="[phrase/morpheme/symbol/word]" type="[inflected/elided/inflectelided/lemma]" xml-
lang="[code langue(-variété-graphie)]" xml:id="form-[identifiant]" (n="[identifiant homographe]")
cert="[current/not_ok/not_ok_var/ok]" value="[collocations + orthographe traitées pour la recherche]">
                    <colloc type="[pre/post]" norm="[forme normalisée]" (orig="[forme originale]" (expand=[forme
pour l'édition numérique]" (subtype="[catégorie détaillée]" value="[collocation traitée pour la
recherche]">collocation usuelle</colloc>*
                        <orth norm="[orthographe normalisée]" (orig="[orthographe originale]" (expand=[orthographe
pour l'édition numérique]" value="[orthographe pour la recherche]">[orthographe de la traduction]</orth>
                        <pron norm="[forme normalisée]" (orig="[forme originale]" (expand=[forme pour l'édition
numérique]" extent="[part/full/pref/suff/poorRhyme/richRhyme]" (notation="[type de
notation]">pronunciation</pron>*
                            <syll norm="[forme normalisée]" (orig="[forme originale]" (expand=[forme pour l'édition
numérique]" extent="[part/full/pref/suff]">découpage en syllabes</syll>*
                                <gramGrp>...</gramGrp>*
                                <bibl (xml-lang="[code langue(-variété-graphie)]">...</bibl>*
                                <lbl type="nature" norm="[forme normalisée]" (orig="[forme originale]" (expand=[forme pour
l'édition numérique]">nature de la forme</lbl>
                                <location>...</location>*
                                <ptr type="audio/img/url/video/other" (subtype="[catégorie]" target="[url]" (targetLang="[code
langue(-variété-graphie)]")></ptr>
                                <usg type="ant/binName/colloc/dom/geo/gram/hint/hyper/lang/plev/register/style/synonym/time"
norm="[forme normalisée]" (orig="[forme originale]" (expand=[forme pour l'édition numérique]" (xml-
lang="[code langue(-variété-graphie)]")>information d'usage</usg>
                                <note rend="visible/invisible" place="above/below">informations diverses</note>*
                            </form>*
                            <gramGrp>
                                <gram type="catEagles" norm="[catégorie grammaticale Eagles]"></gram>?
                                <gram type="catGrace" norm="[catégorie grammaticale Grace]"></gram>?
                                <gram type="catDet" norm="[catégorie détaillée]"></gram>?
                                <gram type="catGen" norm="[catégorie générale]"></gram>?
                                <gram type="catFam" norm="[famille grammaticale]"></gram>?
                                <lbl type="catDet" norm="[catégorie détaillée affichable pleine]" xml:lang="fr" (orig="[forme
originale]" (expand=[forme pour l'édition numérique]">catégorie détaillée abrégée en français</lbl>?
                                    <lbl type="catGen" norm="[catégorie générale affichable pleine]" xml:lang="fr" (orig="[forme
originale]" (expand=[forme pour l'édition numérique]">catégorie générale abrégée en français</lbl>?
                                        <lbl type="catDet" norm="[catégorie détaillée affichable normalisée]" xml:lang="oc" (orig="[forme
originale]" (expand=[forme pour l'édition numérique]">catégorie détaillée abrégée en occitan</lbl>?
                                            <lbl type="catGen" norm="[catégorie générale affichable normalisée]" xml:lang="oc" (orig="[forme
originale]" (expand=[forme pour l'édition numérique]">catégorie générale abrégée en occitan</lbl>?
                                                <pos norm="Part of Speech normalisée"></pos>
                                                <subc norm="Sous-catégorie grammaticale normalisée"></subc>
                                                <per norm="personne normalisée"></per>?
                                                <gen norm="genre normalisé"></gen>?
                                                <number norm="nombre normalisé"></number>?
                                                <tns norm="temps normalisé"></tns>?
                                                <mood norm="mode normalisé"></mood>?
                                                <case norm="cas normalisé pour les pronoms"></case>?

```

```

<gram type="polarity" norm="[polarité normalisée pour l'impératif]"></gram>?
<iType type="accord" norm="[forme normalisée]" (orig="[forme originale]") (expand="[forme pour
l'édition numérique]") (xml:lang="[code langue de l'IANA]")>variable/invariable</iType>*
<iType type="altvb" norm="[forme normalisée]" (orig="[forme originale]") (expand="[forme pour
l'édition numérique]") (xml:lang="[code langue de l'IANA]")>type d'alternance verbale</iType>*
<iType type="inchoative" norm="[forme normalisée]" (orig="[forme originale]") (expand="[forme
pour l'édition numérique]") (xml:lang="[code langue de l'IANA]")>verbe inchoatif ou non inchoatif</iType>*
<iType type="paradigm" norm="[forme normalisée]" (subtype="[sous-type]") (orig="[forme
originale]") (expand="[forme pour l'édition numérique]") (xml:lang="[code langue de l'IANA]")>modèle de
conjugaison</iType>*
<note rend="visible/invisible" place="above/below">informations diverses</note>*
</gramGrp>*
<bibl (xml:lang="[code langue(-variété-graphie)]")>
<author (key="forme normalisée")>
<choice>
<reg>auteur forme à afficher</reg>
<orig>auteur forme originale</orig>?
<expan>auteur forme à afficher édition numérique</expan>?
</choice>
</author>*
<date when="année sur 4 chiffres">
<choice>
<reg>date forme à afficher</reg>
<orig>date forme originale</orig>?
<expan>date forme à afficher édition numérique</expan>?
</choice>
</date>*
<title type="alt/desc/main/short/sub" level="a/m/j/s/u" (key="forme normalisée")>
<choice>
<reg>titre forme à afficher</reg>
<orig>titre forme originale</orig>?
<expan>titre forme à afficher édition numérique</expan>?
</choice>
</title>*
<edition>
<choice>
<reg>numéro d'édition forme à afficher</reg>
<orig>numéro d'édition forme originale</orig>?
<expan>numéro d'édition forme à afficher édition numérique</expan>?
</choice>
</edition>*
<publisher (key="forme normalisée")>
<choice>
<reg>éditeur forme à afficher</reg>
<orig>éditeur forme originale</orig>?
<expan>éditeur forme à afficher édition numérique</expan>?
</choice>
</publisher>*
<pubPlace (key="forme normalisée")>
<choice>
<reg>lieu d'édition forme à afficher</reg>
<orig>lieu d'édition forme originale</orig>?
<expan>lieu d'édition forme à afficher édition numérique</expan>?
</choice>
</pubPlace>*
<region type="type de zone géographique en anglais" (key="forme normalisée")>
<choice>
<reg>région forme à afficher</reg>
<orig>région forme originale</orig>?
<expan>région forme à afficher édition numérique</expan>?

```

```

    </choice>
</region>*

<country key="code ISO">
    <choice>
        <reg>pays forme à afficher</reg>
        <orig>pays forme originale</orig>?
        <expan>pays forme à afficher édition numérique</expan>?
    </choice>
</country>*

<editor role="fonction en anglais" (key="forme normalisée")>
    <choice>
        <reg>responsable forme à afficher</reg>
        <orig>responsable forme originale</orig>?
        <expan>responsable forme à afficher édition numérique</expan>?
    </choice>
</editor*>

<idno type="ISBN/ISSN/DOI/QID/URI/VIAF/ESTC/OCLC">identifiant numérique</idno>*
<biblScope unit="chapter/column/issue/line/page/part/volum">
    <choice>
        <reg>partie forme à afficher</reg>
        <orig>partie forme originale</orig>?
        <expan>partie forme à afficher édition numérique</expan>?
    </choice>
</biblScope*>

<ptr type="audio/img/url/video/other" subtype="catégorie" target="url" (targetLang="code
langue(-variété-graphie)")></ptr>
    <note rend="visible/invisible" place="above/below">informations diverses</note>*
</bibl*>

<lbl type="nature" norm="forme normalisée" (orig="forme originale") (expand="forme pour
l'édition numérique")>nature de la forme</lbl>
    <location>
        <address>
            <postCode>Code postal</postCode>
        </address>?
        <settlement key="code commune" type="city"></settlement>?
        <country key="code ISO">
            <choice>
                <reg>pays forme à afficher</reg>
                <orig>pays forme originale</orig>?
                <expan>pays forme à afficher édition numérique</expan>?
            </choice>
        </country>?
        <region type="type de zone géographique en anglais" (key="forme normalisée")>
            <choice>
                <reg>région forme à afficher</reg>
                <orig>région forme originale</orig>?
                <expan>région forme à afficher édition numérique</expan>?
            </choice>
        </region*>
        <geo>latitude longitude</geo>?
    </location*>
    <ptr type="audio/img/url/video/other" subtype="catégorie" target="url" (targetLang="code
langue(-variété-graphie)")></ptr>*
        <usg type="ant/binName/colloc/dom/geo/gram/hint/hyper/lang/plev/register/style/synonym/time">
            norm="forme normalisée" (orig="forme originale") (expand="forme pour l'édition numérique") (xml-
            lang="code langue(-variété-graphie)")>information d'usage</usg>*
            <note rend="visible/invisible" place="above/below">informations diverses</note>*
        </form*>?
        <person>
            <personName (xml:lang="code langue de l'IANA")>

```

```

<forename sort="[nombre]">prénom</forename>*
<surname sort="[nombre]" type="[birth|married|paternal|maternal]">nom de famille</surname>*
<addName sort="[nombre]" type="[honorific|epithet|nick|stagename]">nom
aditionnel</forename>*
<roleName sort="[nombre]">titre, rôle</roleName>*
<genName sort="[nombre]">nom de génération</genName>*
<nameLink sort="[nombre]">lien entre deux noms</nameLink>*
<forename sort="[nombre]">prénom</forename>*
<choice>
  <reg>nom forme à afficher</reg>
  <orig>nom forme originale</orig>?
  <expan>nom forme à afficher édition numérique</expan>?
</choice>
</personName>+
<birth>
  <date when="année sur 4 chiffres">...</date>*
  <location>...</location>*
</birth>?
<death>
  <date when="année sur 4 chiffres">...</date>*
  <location>...</location>*
</death>?
<gender value="[masculine|feminine|other]">
  <choice>
    <reg>genre forme à afficher</reg>
    <orig>genre forme originale</orig>?
    <expan>genre forme à afficher édition numérique</expan>?
  </choice>
</gender>?
<occupation (from="[date]") (to="[date]")>
  <roleName>
    <choice>
      <reg>rôle forme à afficher</reg>
      <orig>rôle forme originale</orig>?
      <expan>rôle forme à afficher édition numérique</expan>?
    </choice>
  </roleName>?
  <orgName>
    <choice>
      <reg>organisation forme à afficher</reg>
      <orig>organisation forme originale</orig>?
      <expan>organisation forme à afficher édition numérique</expan>?
    </choice>
  </orgName>?
  </occupation>*
<residence (from="[date]") (to="[date]")>
  <location>...</location>
</residence>*
<nationality key="[code ISO]">
  <choice>
    <reg>nationalité forme à afficher</reg>
    <orig>nationalité forme originale</orig>?
    <expan>nationalité forme à afficher édition numérique</expan>?
  </choice>
</nationality>?
<langKnowledge>
  <langKnown level="[native...]" tag="[code langue de l'IANA]">
    <choice>
      <reg>langue forme à afficher</reg>
      <orig>langue forme originale</orig>?

```

```

        <expan>langue forme à afficher édition numérique</expan>?
    </choice>
</langKnown>+
</langKnowledge>?
<idno type="*[type d'identifiant]">identifiant numérique</idno>*
<note rend="*[visible/invisible]" place="*[above/below]">informations diverses</note>*
<ptr type="*[audio/img/url/video/other]" subtype="*[catégorie]" target="*[url]" (targetLang="*[code
langue(-variété-graphie)]")></ptr>*
</person>*
<etym corresp="#*[form/mentioned]-[xml:id d'une forme ou d'un étymon]">
    <lang norm="*[code iso de langue]" (orig="*[forme originale]" (expand="*[forme pour l'édition
numérique]"))>langue d'origine</lang>?
    <mentioned xml:id="*[mentioned-[nombre]">
        <choice>
            <reg>étymon forme à afficher</reg>
            <orig>étymon forme originale</orig>?
            <expan>étymon forme à afficher édition numérique</expan>?
        </choice>
    </mentioned>*
    <gloss (corresp="#*[mentioned-[xml:id d'un mentioned]">
type="*[definition/transcription/translation/type/divers]">
    <choice>
        <reg>glose forme à afficher</reg>
        <orig>glose forme originale</orig>?
        <expan>glose forme à afficher édition numérique</expan>?
    </choice>
    </gloss>*
    <usg type="*[ant/binName/colloc/dom/geo/gram/hint/hyper/lang/plev/register/style/synonym/time]">
norm="*[forme normalisée]" (orig="*[forme originale]" (expand="*[forme pour l'édition numérique]"))>information
d'usage</usg>*
    <bibl>...</bibl>*
    <note rend="*[visible/invisible]" place="*[above/below]">informations diverses</note>*
</etym>*
<sense level="*[nombre]" xml:id="*[sense-[nombre]" (n="*[numérotation à afficher]">
    <usg type="*[ant/binName/colloc/dom/geo/gram/hint/hyper/lang/plev/register/style/synonym/time]">
norm="*[forme normalisée]" (orig="*[forme originale]" (expand="*[forme pour l'édition numérique]")) (xml-
lang="*[code langue(-variété-graphie)]")>information d'usage</usg>*
    <def xml:id="*[def-[nombre]" xml:lang="*[code langue(-variété)(-graphie]" norm="*[forme normalisée]">
(orig="*[forme originale]" (expand="*[forme pour l'édition numérique]"))>définition</def>*
    <cit type="*[translation]">
        <form subtype="*[phrase/morpheme/symbol/word]" type="*[inflected/elided/inflectelided/lemma]">
xml-lang="*[code langue(-variété-graphie)]" xml:id="*[form-[identifiant]" (n="*[identifiant homographe]">
cert="*[current/not_ok/not_ok_var(ok]" value="*[collocations + orthographe traitées pour la
recherche]">...</form>*
        <person>...</person>*
        <etym corresp="#*[form/mentioned]-[xml:id d'une forme ou d'un étymon]">...</etym>*
        <usg type="*[ant/binName/colloc/dom/geo/gram/hint/hyper/lang/plev/register/style/synonym/time]">
norm="*[forme normalisée]" (orig="*[forme originale]" (expand="*[forme pour l'édition numérique]"))>information
d'usage</usg>*
        <note rend="*[visible/invisible]" place="*[above/below]">informations diverses</note>*
    </cit>*
    <xr type="*[abbreviations/antonyms/augmentatives/binName/cf/cognate/contractions/demonyms/
deriveds/diminutives/full/hyperonyms/keywords/lemma/sigle/symbol/synonyms/variants]">
        <lbl type="*[nature]" norm="*[label normalisé]" (orig="*[label original]" (expand="*[label pour l'édition
numérique]"))>label à afficher</lbl>
        <cit type="*[reference]">
            <form subtype="*[phrase/morpheme/symbol/word]" type="*[inflected/elided/inflectelided/lemma]">
xml-lang="*[code langue(-variété-graphie)]" xml:id="*[form-[identifiant]" (n="*[identifiant homographe]">
cert="*[current/not_ok/not_ok_var(ok]" value="*[collocations + orthographe traitées pour la
recherche]">...</form>*
```

```

</cit>*
<person>...</person>*
<bibl>...</bibl>*
<note rend="visible/invisible" place="above/below">informations diverses</note>*
</xr>
<re xml:id="re-[identifiant de sous-entrée"
type="attestation/citation/collocation/compound/demonym/derived/example/expression/interjection/locution/
onomatopeia/proper/pronominal/proverb/unidentified">
    <lbl type="subentry" norm="forme normalisée" (orig="forme originale") (expand="forme pour
l'édition numérique")>label de la sous-entrée</lbl>?
    <form subtype="phrase/morpheme/symbol/word" type="inflected/elided/inflectelided/lemma">
        <!--[code langue(-variété-graphie)]--> xml:id="form-[identifiant]" (n="identifiant homographe")
        cert="current/not_ok/not_ok_var(ok)" value="collocations + orthographe traitées pour la
        recherche">...</form>*
        <person>...</person>*
        <etym corresp="#[form/mentioned]-[xml:id d'une forme ou d'un étymon]">...</etym>*
        <sense level="nombre" xml:id="sense-[nombre]" (n="numérotation à afficher")>...</sense>*
        <note rend="visible/invisible" place="above/below">informations diverses</note>*
        <ptr type="audio/img/url/video/other" subtype="catégorie" target="url" (targetLang="code
langue(-variété-graphie)")></ptr>*
    </re>
    <note rend="visible/invisible" place="above/below">informations diverses</note>*
    <ptr type="audio/img/url/video/other" subtype="catégorie" target="url" (targetLang="code
langue(-variété-graphie)")></ptr>*
    </sense>*
    <xr type="abbreviations/antonyms/augmentatives/binName/cf/cognate/contractions/demonyms/
deriveds/diminutives/full/hyperonyms/keywords/lemma/sigle/symbol/synonyms/variants">...</xr>
    <re xml:id="re-[identifiant de sous-entrée"
type="attestation/citation/collocation/compound/demonym/derived/example/expression/interjection/locution/
onomatopeia/proper/pronominal/proverb/unidentified">...</re>
    <note rend="visible/invisible" place="above/below">informations diverses</note>*
    <ptr type="audio/img/url/video/other" subtype="catégorie" target="url" (targetLang="code
langue(-variété-graphie)")></ptr>*
    </entry>+
</body>
</text>
</xml>

```