

DOSSIÈR DE PREMSA



revirada

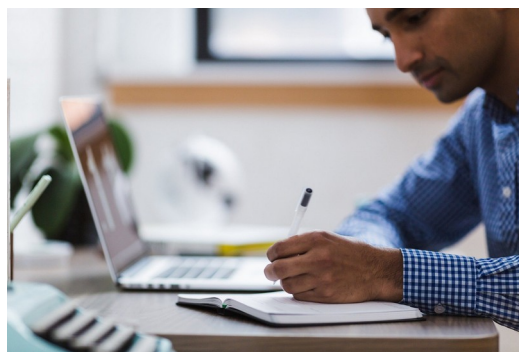
Traductor automatic occitan

Revirada qu'ei lo navèth traductor automatic occitan-francés e francés-occitan deu Congrès entà las varietats gascona e lengadociana.

Revirada qu'ei lo traductor automatic occitan de referéncia deu Congrès. Qu'estó bastit a partir de las melhoras ressorsas de la platafòrma locongres.org, dab **desvolopaments innovants** entà seleccionar los melhors parelhs de traduccion.

Un tribalh deus bèths qu'estó miat autat entà **respectar la diversitat** lingüistica infradialectau com entà **produir ua lenga modèrna, coèrenta e de qualitat**.

Revirada que permet de tradusir tèxtes en copiar-pegar, mes tanben **documents** en conservar la mesa en fòrma o **sites web** « a la volada » a partir de l'URL. Qu'existeishen tanben **aplicacions** entaus telefonets e entaus navigators. Entaus **desvolopaires**, un plug-in e ua API que permeten d'integrar traduccions de qualitat dens los sites web e las aplicacions.



Accedir au traductor automatic en linha *Revirada* : revirada.eu

Dossièr de premsa numeric (dab videos) a l'adreça premsa.locongres.com/revirada.

Dab lo sostien de



Partenaris tecnicos



/ Revirada, quauquas chifras

170 000 mots occitans dont 100 000 gessits deu Congrès	1 milion fòrmas flechidas occitanas mots conjugats o declinats	125 000 mots francés dont 26 000 gessits deu Congrès	540 000 fòrmas flechidas francesas mots conjugats o declinats
170 000 parelhs de traduccion francés-occitan e/o occitan-francés	400 modèles de conjugason en occitan gascon e lengadocian	35 000 terminasons verbaus qui representan la diversitat de l'occitan	
1 100 paradigmes de declinason noms, adjectius, pronoms...	1 600 règlas de seleccion lexicau entà causir la bona revirada		

/ Traductor occitan de referéncia

Revirada qu'ei lo traductor automatic deu Congrès permanent de la lenga occitana. Qu'ei bastit sus la basa deu motor *Open source Apertium* e **qu'intègra las melhoras ressorsas** de la platafòrma occitana de referéncia, locongres.org : 8 dictionaris francés-occitan e occitan-francés, la totalitat de las conjugasons, atau com las basas toponimicas e terminologicas.



Qu'estó necessària mei d'ua annada de tribalh a la nosta equipa entà crotzar e armonizar las dadas deus diferents obratges. **Desvolopaments innovants** que permeten de seleccionar los melhors parelhs de traduccion entà un resultat optimau.



Lo traductor que cobreish **l'ensemble de la diversitat lingüistica infradialectau** dens lo sens occitan-francés : *Revirada*, en efèit, que reconeish mei d'un milion de fòrmas flechidas (mots conjugats e declinats) representativas de la diversitat deus parlars deus parçans gascon e lengadocian.

Dens lo sens francés-occitan, los esfòrç que portèn sus **la produccion d'ua lenga modèrna, coerenta e de qualitat**. Mantun centenat de règlas de seleccion lexicau e algoritmes de pòst-tractament avançats que permeten de garantir ua lenga mei autencica possible.

Aquera purmèra version, limitada peu moment a l'occitan lengadocian e a l'occitan gascon, qu'a per tòca d'enriquesís e de melhorà's. L'utís que serà **regularament actualizat** a mesura de las tornas deus sons utilizators.



/ Las funcionalitats de *Revirada*

Formulari de traduccion



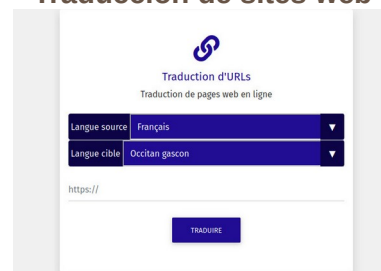
Picar o copiar-pegar tèxte dens l'interfaci web de *Revirada*, causir lo sens e clicar sus un boton.

Traduccion de documents



Teledescargar un document tèxte e *Revirada* que'u ve passarà dens la lenga qui voletz en guardant la mesa en fòrma.

Traduccion de sites web



Entrar l'URL deu site a revirar, la lenga d'origina, la lenga de sortida e qu'auratz lo site revirat.

/ Las aplicacions *Revirada*

Entaus telefonets



L'aplicacion *Revirada* suu vòste telefonet : fotografiatz o copiatz-pegatz, reviratz e escotatz !

Entaus navigators



Teledescargatz la barra de navigator *Linguapir* entà cambiar la lenga deus sites qui visitatz.

Webmèstes e desvolopaires



Ajustatz un boton entà tradusir lo vòste site a la volada, tradusitz los vòstes articles mercés au Plug-in Wordpress, utilizatz la nosta API dens los vòstes programmas.

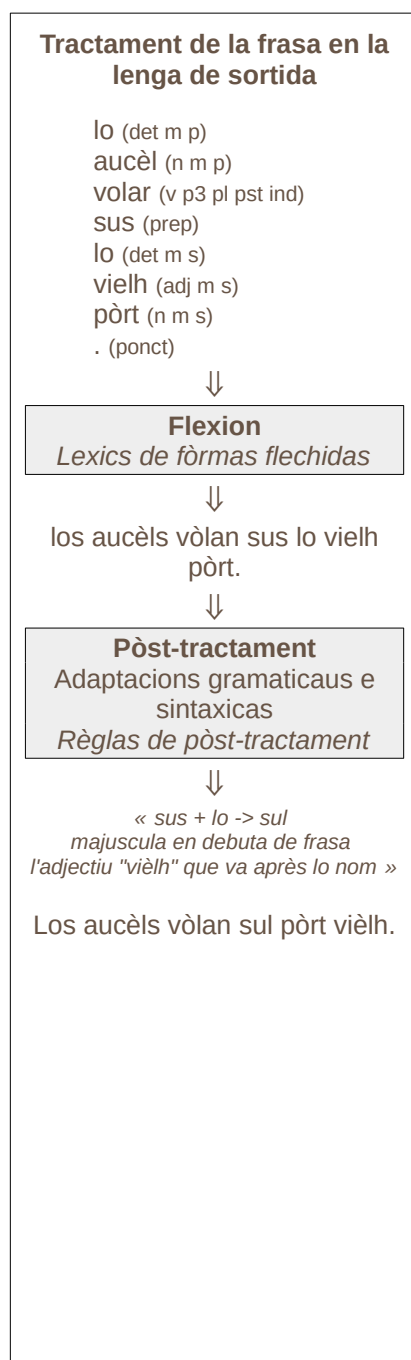
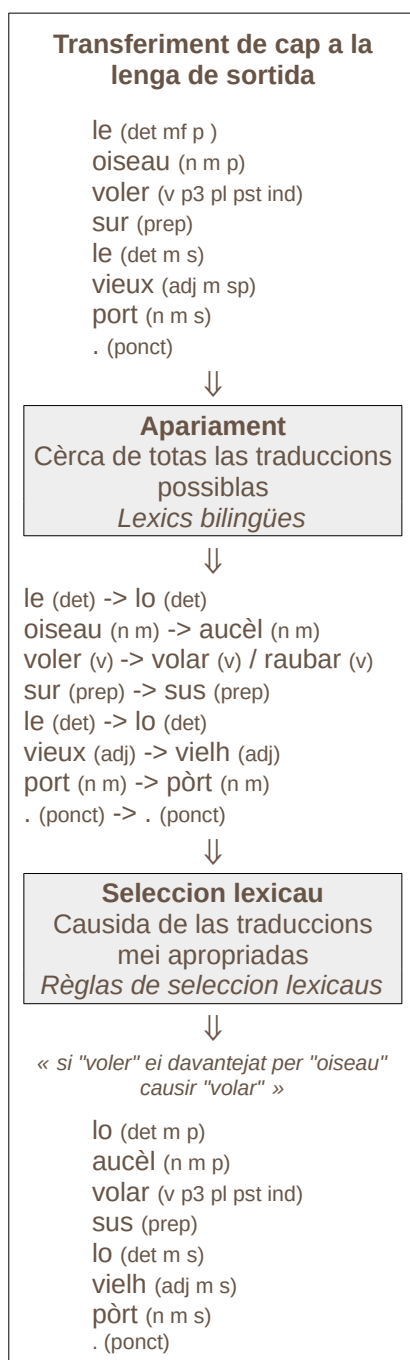
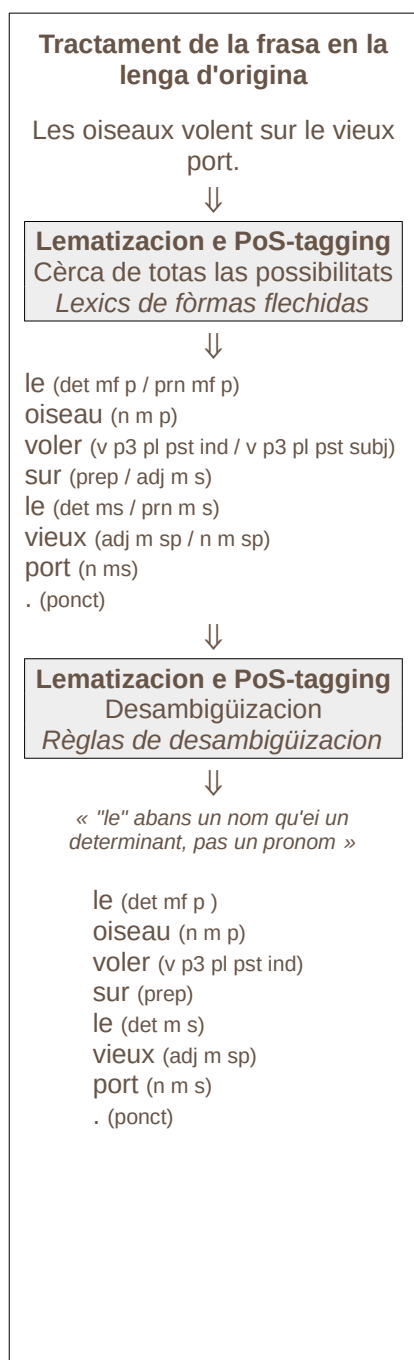
/ *Revirada*, com fonciona ?

Revirada qu'ei bastida a partir deu traductor automatic *Apertium* qui revira en tres etapas :

1. **Tractament de la frasa en la lenga d'origina** : lematizacion (passar a l'infinitiu entaus vèrbes, au masculin singular entaus adjectius...) e analisi morfosintaxica (indicar la categoria gramaticau, lo genre, lo nombre, la persona...).
2. **Transferiment de cap a la lenga de sortida** : entà cada mot de la frasa originau, perpausar la traduccion de qui cau.
3. **Tractament de la frasa en la lenga de sortida** : flexion (tornar conjugar los vèrbes, méter au plurau...) e pòst-tractament (adobar la sintaxi, la gramatica en foncion de las especificitats de la lenga de sortida).

Entad ac har, que's basa sus ressorsas bastidas per umans (au contrari deus traductors qui utilizan l'intelligéncia artificiau, peus quaus lo *machine learning* e basteish automaticament aquera sòrta de ressorsas). Entau cas de *Revirada*, que son estadas creadas **a partir deus lexics deu Congrès** (*dicod'Òc*, *verb'Òc*, *top'Òc*, lexics de fòrmas flechidas), de las coneishenças lexicograficas de l'equipa e de çò qui èra dejà present dens *Apertium*.

- **Lexics de fòrmas flechidas** : que dan modèles de conjugason, de passatge au feminin o au plurau... e listas de mots associadas a un modèle.
- **Règlas de desambigüizacion** : que permeten de causir la bona categoria gramaticau e lo bon lèma quan un mot e'n pòt aver mantun (ex. : « avions » fòrma conjugada deu vèrbe « avoir » o plural deu nom « avion »).
- **Lexics bilingües** : parelhs de mots associats dab la lor traduccion.
- **Règlas de seleccion lexicau** : règlas qui indican quau ei la bona traduccion, si n'i a mantua possibla, en foncion deu contèxte.
- **Règlas de pòst-tractament** : quau ei l'òrdi deus mots, los mots qui's contractan o qui s'eliden, las especificitats idiomáticas...



/ La traduccion automatica e la soa utilitat

En l'estat actua de las tecnologias, la traduccion automatica n'a pas la vocacion de tradusir dab lo medish nivèu qu'un traductor uman. Mes qu'a totun ua utilitat concreta entà mantun contèxte :

- **Comprèner ua lenga desconeishuda** : la traduccion automatica que pòt ajudar quauqu'un a comprèner un tèxte en ua lenga qui ne coneish pas briga. La traduccion qu'aurà dilhèu quauquas errors, l'estile que'n serà desestruc, mes que serà totun sufisent entà que la persona e compenga lo contiengut deu tèxte originau.
- **Avançar lo tribalh deus traductors umans** : si un traductor uman e passa lo son tèxte au traductor automatic, ua part deu tribalh que serà dejà hèita. La persona n'aurà pas mei qu'a corregir lo tèxte generat, meilèu que de s'ac tradusir tot de cap en pè.

Dab **la manca de mejans umans e financèrs** qui an las lengas minorizadas, aquò qu'ei hèra interessant. Per exemple, entà ua estructura qui n'a pas lo vagar, que's pòt adara har lo site web sonque en occitan, e utilizar la traduccion automatica entà'n har lo contienut accessible aus non-occitanofònes. Si, per ua rason o ua auta, e'u hè besonh que la version en francés e sia d'ua bona qualitat, que pòt au mensh utilizar lo traductor automatic entà pretradusir las soas paginas web, e atau panar temps.

/ La gestion de la varietat de la lenga

La **question de la varietat intèrna** de l'occitan que's pausa quasi a cada etapa de la construccion d'un traductor automatic entad aquera lenga. Quau lenga lo traductor e deu rénder ? Com reconéisher lo mei de mots possible ?

Un deus principis d'accion deu Congrès qu'ei lo **respècte de la diversitat** intèrna de la lenga. Lo quasernet de las cargas de *Revirada* que demandava donc que lo traductor entau sens occitan-francés e posca **funcionar quau que sia lo parlar** de l'utilizator (a condicion qui hàcia part de las varietats gascona o lengadociana). Que's miè un tribalh deus bèths entà ensajar d'integrar la diversitat deu vocabulari, de las conjugasons, de las realizacions gramaticaus qui's pòden trobar capvath tota la zòna lingüistica de las duas varietats pertocadas.



Mes aquera diversitat ne devè pas **portar prejudici a la coerència** de la lenga produsida en sortida de la traduccion entau sens francés-occitan. Ne's podèn pas trobar, per exemple, un mot de gascon orientau e un mot de gascon landés en ua medisha frasa. Atau, que's hasó au còp un tribalh de recensament de la diversitat varietau e un **tribalh d'omogeneizacion lingüistica** e de tractament de las variantas.

En qualitat d'organisme de regulacion de la lenga, Lo Congrès qu'avè de **produir ua lenga « normalizada »** dab lo son traductor. Qu'ensagè tanben de seguir lo son principi d'accion qui ei, en mei deu respècte de la diversitat intèrna de la lenga citat ací dessús, **l'ahortiment de la soa unitat pregona**. Tot parièr, qu'ensagè de valorizar los mots comuns e normatius a l'òra de produir los tèxtes en occitan.

Aqueth tribalh que's pòt dejà véder en la version actua de *Revirada*. Que seguirà pendent las annadas a vièner entà que la lenga generada per *Revirada* e sia enquèra mei omogenèa e de qualitat.

/ Linguatec

Lo traductor automatic occitan *Revirada* que's realizè dens l'encastre de **Linguattec**.

Lo **projetè europèu** EFA 227/16 LINGUATEC « Desvolopament de la cooperacion transfronterèra e deu transferiment de coneishenças en tecnologies deu lengatge » qu'a l'objectiu de desvolopar, demostrar e difusar ressorsas navèras, utís e aplicacions lingüistics innovantas entà melhorar lo nivèu de numerizacion de l'**aragonés**, deu **basco** e de l'**occitan**.



Qu'ei portat per un consòrci d'institucions e de centres de recèrca **especializats en lo TAL** (Tractament Automatic de la Lenga) entà las lengas aragonesa, basca e occitana, qui practican la **cooperacion transfronterèra** e lo **transferiment de competéncias**.

Aqueth projetè que permeterà la cooperacion tecnologica entre las lengas entà desvolopar ressorsas navèras, utís lingüistics, atau com aplicacions navèras en linha entà facilitar l'utilizacion de las tres lengas e un accès multilingüe au contiengut peus utilizators finaus.

// Los sòcis de Linguatec



/ Qu'an hargat *Revirada*

// Lo Congrès

Lo Congrès que bastí las ressorsas necessàrias entà har virar lo traductor automatic: lexis, règlas de desambigüizacion, de seleccion lexicau o de pòst-tractament.



Lo Congrès permanent de la lenga occitana qu'ei l'organisme interregionau de regulacion de l'occitan. Qu'òbra en los maines de la lingüistica e deu TAL (tractament automatic de la lenga).

Que produseish utís lingüistics numerics de referéncia (diccionaris, conjugators, correctors ortografics...), aplicacions entau TAL (sintèsi vocau, traduccion automatica...) e aplicacions entaus telefonets (clavèrs predictius...).

Qu'a tanben missions de regulacion lingüistica e de recèrca scientifica aplicada.

Qu'ei l'editor d'un multidiccionari occitan (*dicod'Òc*) qui a cada an mei d'un milion de visitas.

// Elhuyar

Elhuyar qu'a bastit l'interfacci qui utiliza *Revirada* peu mejan d'un formulari o ua API, entà la traduccion de documents o de sites web.



L'objectiu de la fondacion Elhuyar Fundazioa qu'ei d'encoratjar, de promòver e desvolopar la sciéncia e la lenga bascas. En lo maine de las

tecnologias de la lenga e de l'intelligéncia artificial, los sectors de la recèrca suus quaus la Fondacion e tribalha que son, enter autas, la revirada automatica, la creacion automatica de ressorsas lingüísticas, l'extraccion de tèxtes e las tecnologias de la paraula.

Elhuyar que dispausa d'aulhors d'ua experiéncia solida en lo maine deu pilotatge e de la participacion a projèctes europèus de tecnologias deu lengatge. La Fondacion qu'ei sòcia deu Hialat basco de sciéncia e tecnologia.

// Apertium

Apertium qu'ei lo motor de traduccion utilizat per bastir *Revirada*

Apertium qu'ei ua platafòrma *open source* e a gratis de traduccion automatica. Que perpausa un motor de traduccion automatica independent de la lenga, utís entà ajudar a bastir las dadas lingüísticas necessàrias au son foncionament, e dadas déjà disponiblas entà uns parelhs de lenga (com occitan-catalan, occitan-castelhan e occitan-francés).



// Sostiens

Interreg
POCTEFA



/ Entà mei d'informacions

Lo Congrès permanent de la lenga occitana
Ciutat - Creem
5-7 rue de la Fontaine
64000 Pau

premsa.locongres.com/revirada

info@locongres.org

06 45 57 01 07