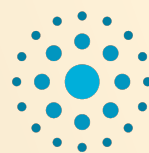


Enquêtes dialectales et collecte de témoignages oraux



Le projet VERA - Vidéo, Ethnotextes et Ressources Associées,
de la linguistique de terrain à la valorisation
du patrimoine culturel immatériel

Pierre-Aurélien GEORGES,
laboratoire BCL



Huma-Num
la TGIR des humanités numériques

CORpus

Langues

Interactions

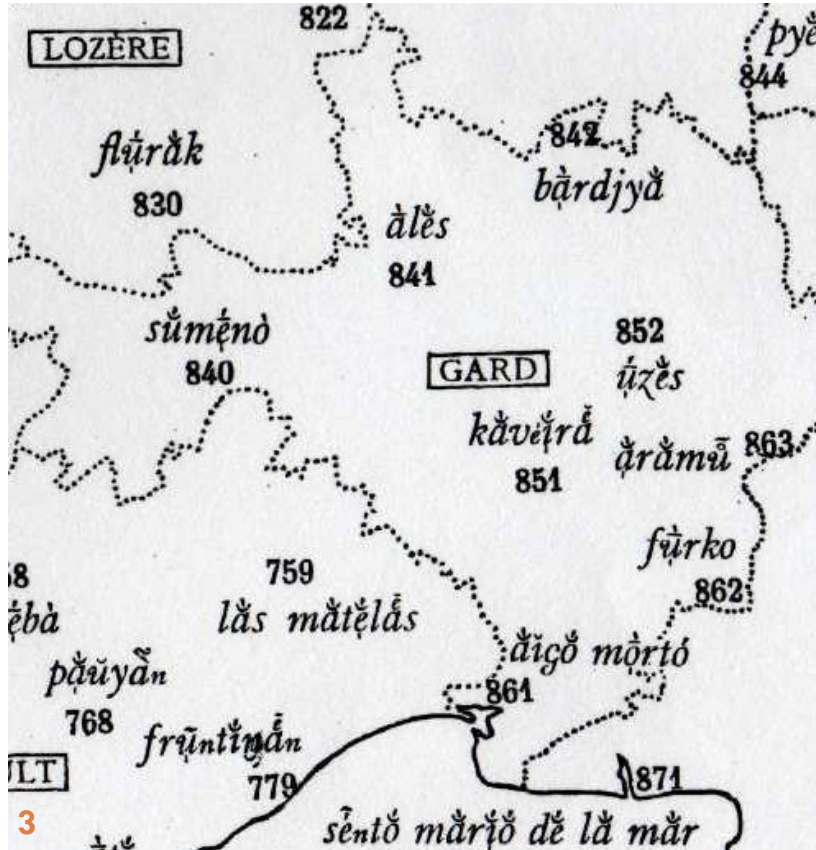




Carte de géographie physique d'Occitanie

© macarel.org

Thesaurus Occitan



- 900+ points d'enquête
- 8000+ questions
- 1,2 million de réponses

<http://thesaurus.unice.fr>



<http://thesaurus.unice.fr>

vera.project.therokuapp.com/search/lexicon?query=nice&term=Nice&type=cities&uid=56343454-eda-64f4-0001

THESOC
La base de données dialectales sur l'Occitanie et ses marges



Nice

Communes

Filtrer résultats


Medias **Lexique (34)**

Thèmes

Thèmes	Notion	Catég.		
Famille Homme	père	[+]		paire (PATER) ALLy 945 ALMC 1607 ALAL 1027 [...] THESOC Q7562
Oiseaux Nature	pinson (<i>Fringilla coelebs</i>)	[+]	Substantif	*alpin °°balarina pinçon (*PINCIO) [...] ALLy 501 ALMC 316 ALAL 418 [...] THESOC Q7719
Oiseaux Nature	 huppe (oiseau) (<i>Upupa epops</i>)	[+]	Substantif	puput (ÜPÜPA) huppe (fr) (ÜPÜPA) °°pupa [...] ALMC 319 ALO 421 ALCe 548 [...] THESOC Q5587
Oiseaux Nature	 roitelet (<i>Regulus regulus</i>)	[+]	Substantif	°°boscarla °°boscarleta [...] ALLy 512 ALMC 316

Q saint-jean

- Saint-Pierre-Saint-Jean
Communes
- bête à saint Jean (fr)
Lemmes
- corne de saint Jean (fr*)
Lemmes
- faire saint Jean (fr)
Lemmes
- feu de la Saint-Jean
Questions



Les spécificités du THESOC

- Oralité (linguistique de terrain)
- Cartographie (toutes les données sont géolocalisées)
- Grande diversité de dialectes
(versus occitan « standardisé » / « langues nationales »)

Pourquoi de nouvelles enquêtes ?

6

- Compléter les données des atlas linguistiques
- Données plus récentes (diachronie)
- Multimédia (audio/vidéo et texte synchronisés)
- Questionnaire élargi : syntaxe, discours libre



Enquête linguistique à Breil-sur-Roya (Alpes-Maritimes), 2008.
Michèle Olivéri & Jean-Claude Ranucci





L'épreuve de chant
du Certificat d'Études



La fabrication du
fromage



L'âge des brebis



Les 5 repas du
paysan
limousin



Le métier de
typographe

1980

2002

CENTRE NATIONAL DE LA RECHERCHE SCIENTIFIQUE
Centre Régional de Publications: Marseille

TRADITION ORALE ET IDENTITE CULTURELLE

problèmes et méthodes

Jean-Claude BOUVIER, Henry-Paul BREMONDY
Philippe JOUTARD, Guy MATHIEU, Jean-Noël PELEN



EDITIONS DU CNRS

« *l'ethnotexte est principalement un discours sur le passé qui est recueilli [...] pendant un entretien, ou une enquête, dont la thématique appartient au discours culturel oral d'une communauté.* »

Pierre Chabert

Ethnotextes : la conscience de l'espace. L'exemple d'un conte, bulletin de l'AFAS n°22, p. 2-8

<https://journals.openedition.org/afas/1469>

2006 – 2009 : HOROYA

Histoire Orale des vallées de la Roya et Bévéra

Resp. Th. Rosso (LAPCOS), 2006-2009.



2010 – 2012 : PHC Procope

Syntactic Microvariation : Clitic Pronouns in Romance Languages

Resp. M. Oliviéri (BCL) & G. Kaiser (Universität Konstanz, Allemagne)



2012 – 2015 : DADDIPRO

Dialectal, acquisitional, and diachronic data and investigations on subject pronouns in Gallo-Romance

Resp. M. Oliviéri (BCL) & G. Kaiser (Universität Konstanz, Allemagne)



2013 – 2015 : SyMiLa

Syntactic Microvariation in the Romance Languages of France

Resp. P. Sauzet (Université Toulouse Jean Jaurès) & D. Sportiche (Paris)



2006 – 2017



100+ enquêtes de terrain



75 localités

(France méridionale & Italie)



120+ locuteurs interviewés



1200+ questions posées

(5 questionnaires de phrases)



60 h de vidéo inédites à traiter

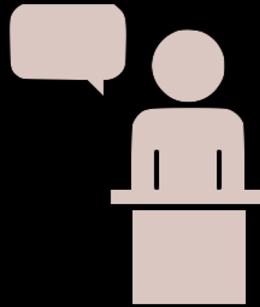


58 ethnotextes récoltés

3 publics identifiés



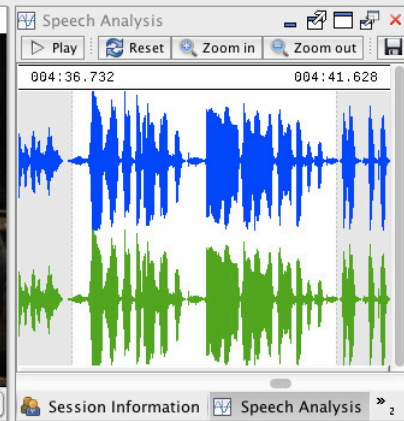
Grand Public



**Apprenants
Enseignants
Patoisants**



**Linguistes
Ethnologues
...**



Record Data

65 Speaker: Pascal Martini Exclude from searches

Segment [004:36.732-004:41.628]

IPA Target [e 'ʃi ke i pɔ: fa: aɪ 'brusu: ! me l eʒ ɛ tra'vaju ke nu 'fasu tʃy]

Graphie [E sci che i pòo fâa ař brussu ! <i lang="fr">Mais </i> l'ei 'n travayu che nu fassu ciü.]

Traduction [Et bien sûr qu'on peut faire le brousse ! Mais c'est un travail que je ne fais plus.]

Notes [Le "brousse" désigne un fromage de lactosérum originaire de Provence, à la pâte blanche et à la texture grumeleuse.]

traduction en Italien [Ma certo che si puo' fare la brossa! Ma è un lavoro che non faccio più.]

id: 6f6e49b5-48d8-4d51-8bbf-d9248642ae59 Tier: Group: Character:





Record List :

Segment	Speaker	IPA Actual	Traduction	Graphie	IPA Target	traduction en Italien
002:23.202...	Jean-Cla...	[Je vais vous demander un secret : comment vous faites le fromage ?]				Vi chiederò 'un segret...
002:26.914...	Pascal M...	[]	Ah, pour faire le fromage, c'est un secret que je ne le dis à personne.	Ah, pef fâa ař fru...	[a: pœ fa: aɪ fr...	Ah, per fare il formag...
002:34.066...	Jean-Cla...	[Expliquez-moi un peu...]			[]	Mi spieghi un po'...
002:35.856...	Pascal M...	[]	Ah, je ne peux, non... Ah non, non, non, non, non...	Ah, nu pösciü, nu...	[a: nu 'pœʃu, nu...	Ah, non posso, no...n...
002:39.942...	Pascal M...	[]	et après les autres me prennent le brevet, hein !	E dopu i autri me...	[e 'dɔpu i 'jawtri ...	e dopo gli altri mi pr...
002:46.048...	Jean-Cla...	[Ils n'ont plus de vaches ! Il faut faire chauffer le lait ?]			[]	Ma no, non hanno pi...
002:51.807...	Pascal M...	[]	Eh ! Il faut faire chauffer le lait à environ 35°C,	Eh ! Venta fâa sca...	[e: 'vɛta fa: ʃka...	Eh, bisogna far scalda...
002:56.760...	Pascal M...	[]	et puis il faut y mettre la présure,	e pöi venta meteř...	[e pœʒ 'vɛta mete...	e poi bisogna metter...
002:59.426...	Pascal M...	[]	puis le laisser reposer, couvert, quarante minutes, à peu près quarante minutes,	pöi lasciařlù paus...	[pœʒ laʃař lu pa...	per lasciarlo riposare...
003:06.261...	Pascal M...	[]	puis ensuite il faut le frapper avec le tranche-caillé ;	pöi dopu venta sb...	[pœʒ 'dɔpu 'vɛta ...	poi in seguito bisogna...
003:11.703...	Pascal M...	[]	puis, il doit être rassemblé avec les mains, puis le prendre avec l'étamine,	pöi ař va recampa...	[pœʒ aɪ va rekäp...	poi, va ricomposto co...
003:18.213...	Pascal M...	[]	et puis le mettre à égoutter, le mettre à... le mettre à égoutter.	e pöi u mete a res...	[e pœʒ u 'mete a r...	e poi metterlo a scola...
003:25.292...	Jean-Cla...	[Et après on le met dans la forme ?]			[]	E dopo lo si mette nel...
003:29.286...	Pascal M...	[]	Eh... moi je ne le mets pas dans la forme, car je le fais carré, avec les mains.	Eh... mi nu u met...	[e: mi nu u 'met...	Eh... io non lo metto...
003:36.024...	Jean-Cla...	[Et après il est fait au bout de combien de temps ? Il faut le saler ?]			[]	E dopo, in quanto te...
003:39.290...	Pascal M...	[]	Ah ! Il faut le saler... le soir il faut le saler d'un côté...	Ah ! Venta sařařlù...	[a: 'vɛta saɪaɪ l...	Ah, bisogna salarö, la...
003:43.271...	Pascal M...	[]	Le lendemain matin, il faut le tourner et le saler de l'autre côté...	Ř'endeman de ma...	[ɹɛdəmä de mař, ...	L'indomani di mattina...
003:49.663...	Pascal M...	[]	Et puis il doit être essuyé, il doit être tourné, et de temps en temps il faut s'occuper, ...	e pöi ař va sciügä...	[e pœʒ aɪ va ʃyğ...	Poi va asciugato, va gi...
003:56.988...	Jean-Cla...	[Et combien de temps il faut le garder avant de le vendre ?]			[]	Ma quanto tempo bis...
004:00.347...	Pascal M...	[]	Et ça dépend : il y en a qui les vendent même frais, à partir de...	E ça depende : ař ...	[e sa de'pɛde a...	Eh... dipende: ci sono ...

Save as CSV...



Sous-titres au choix :

- français  
- occitan 
- phonétique API 
- etc.

Mon père s'était échappé des allemands ; on était là-bas dans la montagne.



1:13 / 1:57



BASES
CORPUS
LANGAGE



ANR-15-CE0007-01





L'hiver 1956 : le plus froid du XX^e siècle.

L'hiver 1956.

L'hiver 1956 est marqué, dans l'hémisphère Nord, par une vague de froid sans précédent. Historiquement, il rappelle le Grand Hiver de 1709 qui voit le vin geler dans les coupes de Versailles tout comme l'encre des encriers de l'école de Pierre Lapomroulle à Sorges dans le Périgord en 1956. Cet épisode n'est pas isolé puisque, de 1999 à 2010, le Grand Hiver intense touche la France. Il en sera de même en 1963, en 1985 et 2009-2010. Cependant, les records de température font le mois le plus froid du XXI^e siècle.

Les bombardements à Nice Seconde Guerre mondiale.

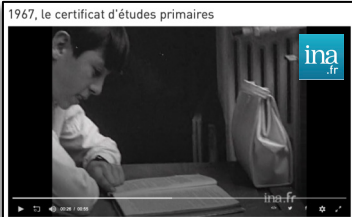
Entre 1943 et 1944, plusieurs séries de bombardements sont effectués sur la ville de Nice. Le 1 et le 3 mai, les premières bombes atteignent Nice. Une nouvelle série de bombardements a lieu le 26 mai 1944 et détruit l'Arénas. Par la suite, une dizaine de bombardements touchent les ponts du Var, l'Hippodrome, le Port et le Vieux-Nice.

Le bombardement du 26 mai 1944 compte parmi les plus importants de la Seconde Guerre mondiale. De nombreux dégâts dans la ville. Dans le cadre du Transportation Plan, la destruction des communications dans le Sud de la France doit préparer le défilé de l'armée alliée. L'objectif est de détruire les infrastructures ferroviaires notamment la gare de Saint-Roch. Pour cela, l'armée alliée adopte la méthode « américaine » : le choix de largage à haute altitude présente des avantages certains : moins de victimes civiles. Cette opération militaire vise la gare de Saint-Roch et la gare de Riquier, de Pasteur et de République. Le bilan est d'environ 350 victimes.

Plusieurs plaques commémoratives rappellent cet événement au 100, au 102, au 104, au 106, au 108, au 110, au 112, au 114, au 116, au 118, au 120, au 122, au 124, au 126, au 128, au 130, au 132, au 134, au 136, au 138, au 140, au 142, au 144, au 146, au 148, au 150, au 152, au 154, au 156, au 158, au 160, au 162, au 164, au 166, au 168, au 170, au 172, au 174, au 176, au 178, au 180, au 182, au 184, au 186, au 188, au 190, au 192, au 194, au 196, au 198, au 200, au 202, au 204, au 206, au 208, au 210, au 212, au 214, au 216, au 218, au 220, au 222, au 224, au 226, au 228, au 230, au 232, au 234, au 236, au 238, au 240, au 242, au 244, au 246, au 248, au 250, au 252, au 254, au 256, au 258, au 260, au 262, au 264, au 266, au 268, au 270, au 272, au 274, au 276, au 278, au 280, au 282, au 284, au 286, au 288, au 290, au 292, au 294, au 296, au 298, au 300, au 302, au 304, au 306, au 308, au 310, au 312, au 314, au 316, au 318, au 320, au 322, au 324, au 326, au 328, au 330, au 332, au 334, au 336, au 338, au 340, au 342, au 344, au 346, au 348, au 350, au 352, au 354, au 356, au 358, au 360, au 362, au 364, au 366, au 368, au 370, au 372, au 374, au 376, au 378, au 380, au 382, au 384, au 386, au 388, au 390, au 392, au 394, au 396, au 398, au 400, au 402, au 404, au 406, au 408, au 410, au 412, au 414, au 416, au 418, au 420, au 422, au 424, au 426, au 428, au 430, au 432, au 434, au 436, au 438, au 440, au 442, au 444, au 446, au 448, au 450, au 452, au 454, au 456, au 458, au 460, au 462, au 464, au 466, au 468, au 470, au 472, au 474, au 476, au 478, au 480, au 482, au 484, au 486, au 488, au 490, au 492, au 494, au 496, au 498, au 500, au 502, au 504, au 506, au 508, au 510, au 512, au 514, au 516, au 518, au 520, au 522, au 524, au 526, au 528, au 530, au 532, au 534, au 536, au 538, au 540, au 542, au 544, au 546, au 548, au 550, au 552, au 554, au 556, au 558, au 560, au 562, au 564, au 566, au 568, au 570, au 572, au 574, au 576, au 578, au 580, au 582, au 584, au 586, au 588, au 590, au 592, au 594, au 596, au 598, au 600, au 602, au 604, au 606, au 608, au 610, au 612, au 614, au 616, au 618, au 620, au 622, au 624, au 626, au 628, au 630, au 632, au 634, au 636, au 638, au 640, au 642, au 644, au 646, au 648, au 650, au 652, au 654, au 656, au 658, au 660, au 662, au 664, au 666, au 668, au 670, au 672, au 674, au 676, au 678, au 680, au 682, au 684, au 686, au 688, au 690, au 692, au 694, au 696, au 698, au 700, au 702, au 704, au 706, au 708, au 710, au 712, au 714, au 716, au 718, au 720, au 722, au 724, au 726, au 728, au 730, au 732, au 734, au 736, au 738, au 740, au 742, au 744, au 746, au 748, au 750, au 752, au 754, au 756, au 758, au 760, au 762, au 764, au 766, au 768, au 770, au 772, au 774, au 776, au 778, au 780, au 782, au 784, au 786, au 788, au 790, au 792, au 794, au 796, au 798, au 800, au 802, au 804, au 806, au 808, au 810, au 812, au 814, au 816, au 818, au 820, au 822, au 824, au 826, au 828, au 830, au 832, au 834, au 836, au 838, au 840, au 842, au 844, au 846, au 848, au 850, au 852, au 854, au 856, au 858, au 860, au 862, au 864, au 866, au 868, au 870, au 872, au 874, au 876, au 878, au 880, au 882, au 884, au 886, au 888, au 890, au 892, au 894, au 896, au 898, au 900, au 902, au 904, au 906, au 908, au 910, au 912, au 914, au 916, au 918, au 920, au 922, au 924, au 926, au 928, au 930, au 932, au 934, au 936, au 938, au 940, au 942, au 944, au 946, au 948, au 950, au 952, au 954, au 956, au 958, au 960, au 962, au 964, au 966, au 968, au 970, au 972, au 974, au 976, au 978, au 980, au 982, au 984, au 986, au 988, au 990, au 992, au 994, au 996, au 998, au 1000.

Une autre série de bombardement a marqué les esprits des Niciens le 27 août. Un avion, non identifié mais probablement américain, survole la ville et ravage la ville. La nuit du 13 août huit bombes sont ainsi lâchées sur la ville. Les bombardements se poursuivent dans les nuits suivantes et touchent notamment le Mont-Boron, la Rue des Orangers, Massena, le Vieux-Nice et de nombreuses autres zones de la ville.

Le Certificat d'études primaires



Le Certificat d'études primaires (CEPE) est le premier diplôme de l'école primaire. Il est créé en 1882 par le ministre de l'Éducation nationale Napoléon III. Ce diplôme est destiné aux élèves de la classe de sixième. Il est remplacé en 1959 par le Certificat de fin d'études primaires (CFEP).

Le dernier est en fait le plus ancien mode d'évaluation, celui qui a subi la concurrence d'un autre mode d'évaluation, celui des concours cantonaux qui opposent les meilleurs élèves de chaque école.

Le 16 juin 1880, un arrêté ministériel harmonise les modalités d'évaluation du certificat d'étude. Désormais le nombre de cinq fautes à la dictée devient éliminatoire.



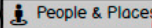
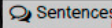
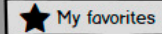
Le 16 juin 1880, un arrêté ministériel harmonise les modalités d'évaluation du certificat d'étude. Désormais le nombre de cinq fautes à la dictée devient éliminatoire.



How to make traditional cheese

St-Martin-du-Var (06670), juin 2014.

Ethnotext



people, place, word, atlas map, ...



Informant(s) : [Pascal Martini](#)

Source : [DADDIPRO](#)

Export text



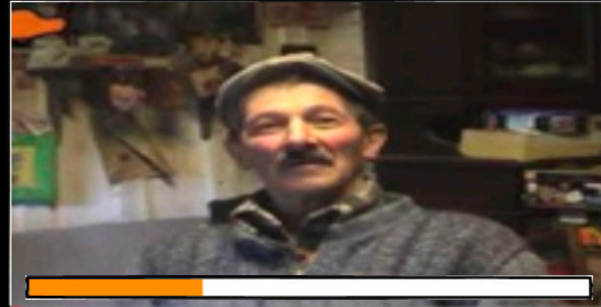
Autoplay



IPA transcriptions



Recording



Historical Background

Les Italiens à Nice

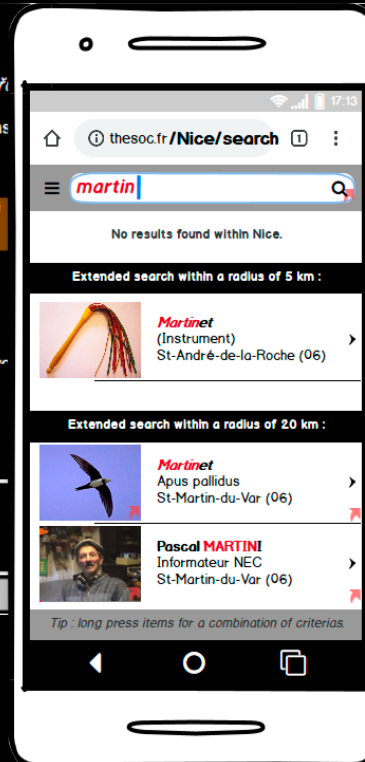
À Partir de l'annexion du Comté de Nice, de nombreux Italiens s'installent à Nice et particulièrement dans la Vieille-Ville. Leur nombre croît jusque dans la première moitié du XXe siècle. En 1975, Paul Caramagna étudie cette communauté présente à Nice en 1921 en s'appuyant sur le recensement de la population. 27% des habitants de Nice sont étrangers et 22,6% sont d'origine italienne dont seulement 48% sont des hommes. Ils habitent majoritairement dans la Vieille-Ville selon un axe Garibaldi-Port. En 1921, ce sont ainsi 12 pêcheurs italiens qui vivent dans le quartier du Malonat. La

User Comments about this ethnotext

F. G. Le su
grand
J.-C. R. En ter
F. G. En ter
De mi
taurel
J.-C. R. Quel i
F. G. Hé... j
Le ve

1. *tofu* désigne le

de... de mon
e c'était le *tofu*
que, en tendas
cu...
egran, l'e'a af
avez-vous ?
ni...
a af me che ur



Thesaurus Occitan (THESOC)

https://thesoc.fr/sentences?ipa#map+facets

THESOC - Thesaurus Occitan
La base de données dialectales sur l'Occitanie et ses marges

Media | Lexicon | Sentences | People & Places

Questions | Productions | Export selection

IPA transcriptions

Ref.	People & Location	Sentence
Question NEC 195	José TRUCHAUD St-Martin-du-Vaï (06) ALF 872 · ALP 34	Quoura eravan jouve, avèn passat de bouòn moumen mé eli. [k' uʁa eʁ' avən dʒ' uve l av' èm pas' at de bwõm mum' èm me' eli] <i>When (we) were young, (we) have spent some good moments with the</i> Sujet nul compl. « de » aux. avoir adverbe
Question NEC 195	Jean-Claude PIERRE Eymoutiers (87) ALF 435 · ALAL 34	Quand nos eron juno, nos passerem de brave temps como eli. [kã nu eʁõ dzyno nu pãʒeʁe] de brav te] kumo j] <i>When we were young, we spent some good time with them.</i> compl. « de » aux. avoir adverbe
Question PAM2 256	José TRUCHAUD St-Martin-du-Vaï (06) ALF 872 · ALP 34	Eh, n'en fan de caua ! <i>Hey, (she_PL) it[PRO] make lot_of things !</i> Sujet nul compl. « de »
Question PAM2 256	Jean-Claude PIERRE Eymoutiers (87) ALF 435 · ALAL 34	[a læ ne] fõ dœ la soza ! <i>Hey, she_PL it[PRO] make lot_of the thing.</i> compl. « de »

Map

Presence of Sujet nul

Eymoutiers (87)
ALF 435 · ALAL 34
Jean-Claude Pierre
« de » | aux. avoir | adverbe

Facets

- Filter by genre
- Filter by questionnaire
- Filter by tags
 - Sujet : Sujet nul | Explétif | Impersonnel
 - Verbe : aux. avoir | aux. être

Users comments about this question

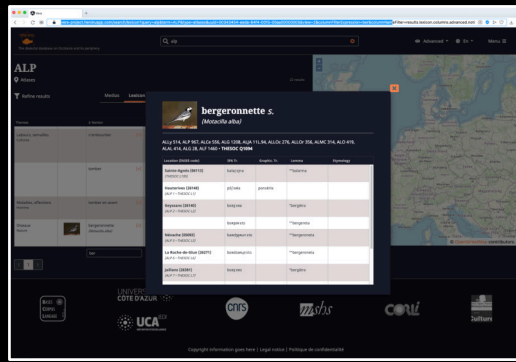


Annotations linguistiques

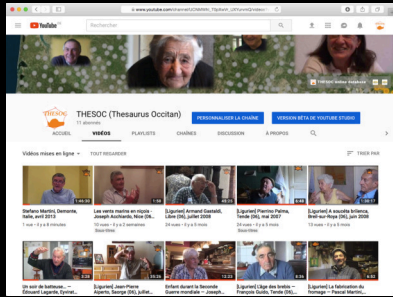
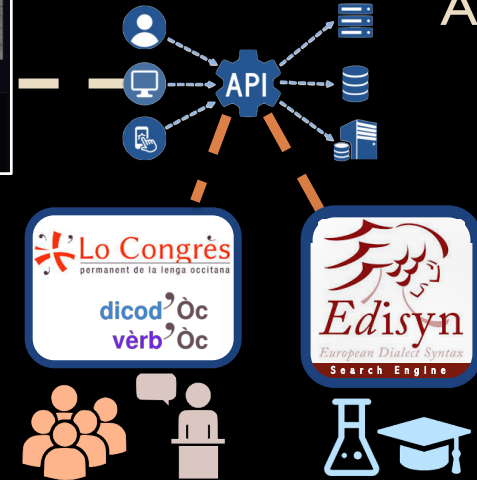
→ cartographie de phénomènes syntaxiques ou lexicaux



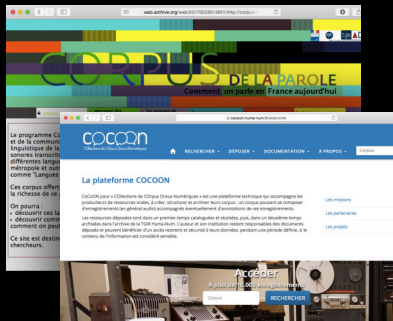
Plateforme web dédiée
cartes, concordancier, ...



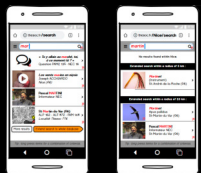
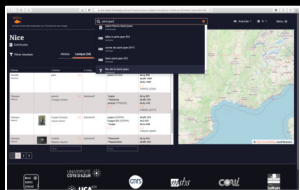
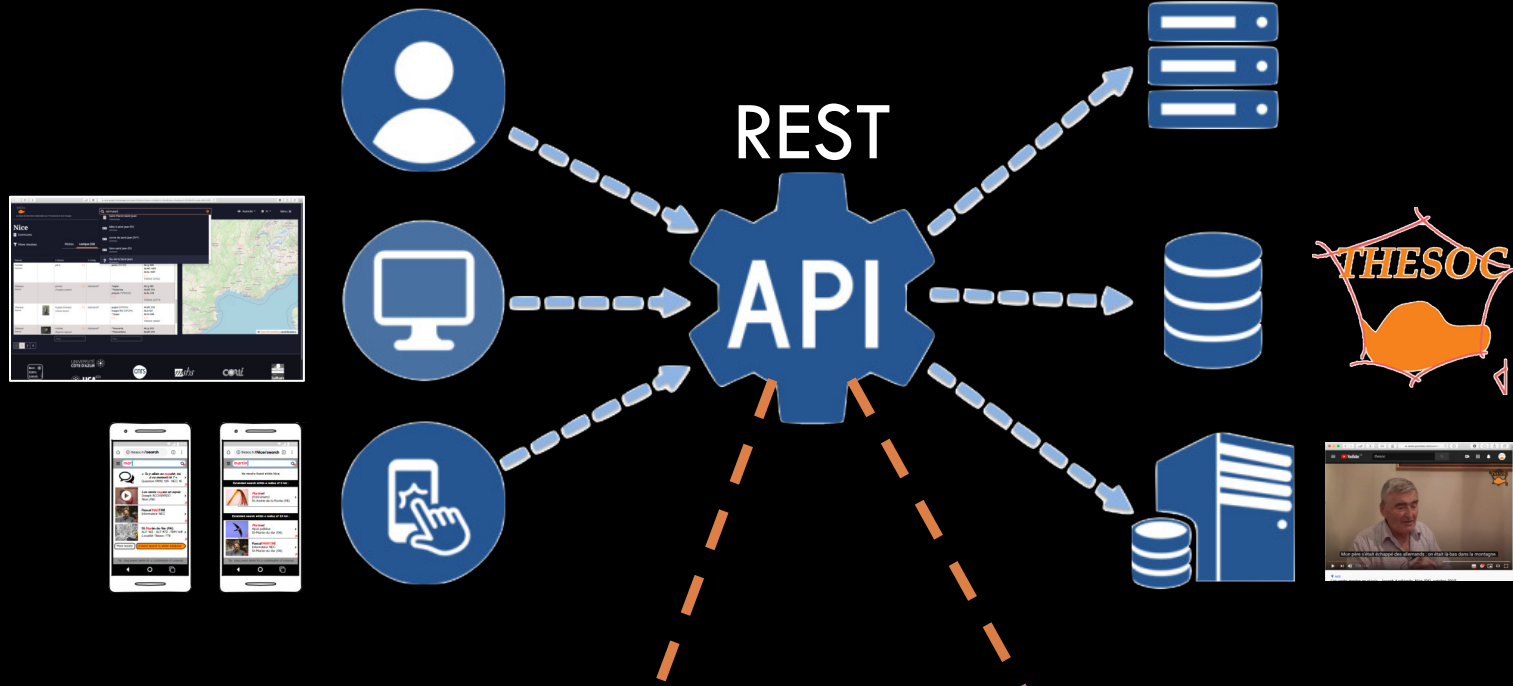
Interopérabilité
API REST



Chaîne YouTube
ethnotextes sous-titrés



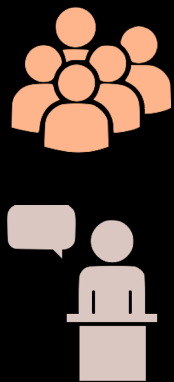
Archivage pérenne
+ Corpus de la parole



Lo Congrès
permanent de la lenga occitana

dicod'Òc
vèrb'Òc

Edisyn
European Dialect Syntax
Search Engine





Formats de fichier standards

Enregistrements en MPEG H264 + .WAV

Ressources en fichiers XML / TEI P5



Licences ouvertes

Vidéos en CC-BY-NC-ND

Tout le reste en **LGPL-LR**



Métadonnées

Dublin Core & OLAC

multiples applications possibles

- **Traitement Automatique du Langage**
(traduction automatique, reconnaissance vocale, etc.)
- **Linguistique** (dialectologues, phonologues, syntacticiens ;
mais aussi communication non verbale, etc.)
- **Autres disciplines scientifiques** (historiens, sociologues, ...)
- **Apprentissage des langues régionales**
- *Big-data / Deep Learning*

Quelques conseils

- Ne pas filtrer les données à priori
(difficile de prévoir toutes les exploitations possibles)
- Témoins s'exprimant dans leur dialecte vernaculaire
(+ naturel et spontané ; patrimoine linguistique intéressant)
- Qualité des enregistrements et des données récoltées
Manipulation du matériel, déroulement de l'enquête
- Penser à récolter les données annexes / métadonnées
Photos, croquis, relevés GPS, etc.
- Autorisations des témoins
(droit à l'image, vie privée, anonymat, etc.)