

Construction du lexique *LGLex* à partir des tables du Lexique-Grammaire des verbes du grec moderne



ΑΡΙΣΤΟΤΕΛΕΙΟ
ΠΑΝΕΠΙΣΤΗΜΙΟ
ΘΕΣΣΑΛΟΝΙΚΗΣ

UNIVERSITÉ
— PARIS-EST



Kyriaki IOANNIDOU
LTTL, Université Aristote de Thessalonique
Grèce

Elsa TOLONE
LIGM, Université Paris-Est
France

&
FaMAF, Universidad Nacional de Córdoba
Argentine

Modèle linguistique



- **Méthodologie du Lexique-Grammaire (Gross, 1975)**
 - approche symbolique – développement de ressources linguistiques
 - inventaire des prédicats divisés en classes selon leurs propriétés définitoires
 - description syntaxico-sémantique des prédicats et de ses compléments
 - compatibilité des données grecques avec celles de l'équipe française
- **Notre but : exploiter nos ressources linguistiques pour l'analyse syntaxique**

Analyse syntaxique



- Logiciel d'analyse syntaxique
- Un ensemble de grammaires/métagrammaires
 - formation des phrases
 - reconnaissance des syntagmes simples et complexes
 - dépendances entre les différents syntagmes dans la phrase
- Ressources lexicales
 - dictionnaires (morphologiques – syntaxiques – sémantiques)
 - *LGLex*

Plan de travail



- Traitement des ressources linguistiques : tables du LG (Ioannidou & Kyriacopoulou, 2010 ; Tolone 2011)
 - collection des tables existantes
 - ✦ permet l'extraction des informations depuis toutes les tables et non pas quelques tables séparément
 - homogénéisation des données
 - ✦ n'était pas assurée à cause de l'existence de plusieurs auteurs. Application séparée de chaque table jusqu'à maintenant grâce à des graphes paramétrés
 - ✦ est nécessaire pour qu'il y ait de la cohésion parmi les tables
 - ✦ est facilitée par la table des classes
 - création d'un seul fichier et transcodage des informations en un format exploitable par des analyseurs syntaxiques
 - ✦ permet l'application de toutes les tables à la fois
 - ✦ permet son exploitation facile (xml/txt ≠ .xls) et son intégration dans divers logiciels pour procéder à une analyse syntaxique

Tables du Lexique-Grammaire traitées



- Verbes transitifs non-locatifs (Voskaki, 2011)
(32GA, 32GD, 32GC, 32GCL, 32GCV, 32GH, 32GNM, 32GPL, 32GRA)

○ Άνοιξα το παράθυρο

32GC : N0 V N1 οὐ N1 := concret

○ J'ai ouvert la fenêtre

- Verbes à complétives (Kyriacopoulou, 2005; Fista, 2007)
(4G, 6G, 9G)

○ Άκουσα ότι φεύγεις

6G : N0 V N1 οὐ N1 := Pcomp

○ J'ai entendu que tu pars

Tables du Lexique-Grammaire traitées



- Verbes transitifs locatifs

(Yannacopoulou 2005 ; Fista 2007)

(38GL, 38GLS, 38GLD, 38GLH, 38GLR)

○ Έβαλε το ποτήρι πάνω στο τραπέζι 38GLS N0 V N1 Loc N2 source

○ Il a mis le verre sur la table

Colonnes non traitées



- Traduction et table correspondante de la table française

N0 =: Nhum	N0 =: N-hum	N0 =: Vh	N1 V	V	N0 V N1	N1 =: disp	N1 =: Nhum	N1 =: N-hum	N1 =: ppv	N1 είναι V:K	Npréd	Vsup =: προκαλώ	Vsup =: προχωρώ σε, προβαίνω σε	VP	N1 VP από N0	N1 VP	Traduction	Table FR
+	+	αφανιστής	-	αφανίζω	+	+	+	+	+	+	αφανισμός	+	+	αφανίζομαι	+	+	ravager	32C
-	+	-	-	βαραθρώνω	+	+	+	+	+	+	βαραθρώση	+	-	βαραθρώνομαι	+	+	détruire	32D
+	+	γκρεμιστής	-	γκρεμάω	+	+	-	+	+	+	γκρέμισμα	+	+	γκρεμιέμαι	+	+	démolir	32D
+	+	γκρεμιστής	-	γκρεμίζω	+	+	-	+	+	+	γκρέμισμα	+	+	γκρεμίζομαι	+	+	démolir	32D
+	+	γκρεμιστής	-	γκρεμίζω	+	+	-	+	+	+	-	-	-	-	-	-	démolir	32D
+	+	γκρεμιστής	-	γκρεμιάω	+	+	-	+	+	+	γκρέμισμα	+	+	γκρεμιέμαι	+	+	démolir	32D
+	+	-	-	εκμηδενίζω	+	+	-	+	+	+	εκμηδένιση	+	+	εκμηδενίζομαι	+	+	annihiler, anéantir	32R2, 32D
+	+	-	-	εξανεμίζω	+	+	-	+	+	+	εξανέμιση	+	+	εξανεμίζομαι	+	+	volatiliser	32D
+	+	-	-	εξαντλώ	+	+	-	+	+	+	εξάντηληση	+	-	εξαντλούμαι	+	+	épuiser	32PL
+	+	-	-	εξαφανίζω	+	+	+	+	+	+	εξαφάνιση	+	-	εξαφανίζομαι	+	+	disparaître, supprimer	35ST, 32D
+	+	εξολοθρευτής	-	εξολοθρεύω	+	+	-	+	+	+	εξολόθρευση	+	+	εξολοθρεύομαι	+	+	anéantir, détruire, exterminer	32D, 32D, 32H
+	+	-	-	εξοντώνω	+	+	+	-	+	+	εξόντωση	+	+	εξοντώνομαι	+	+	anéantir, détruire, exterminer	32D, 32D, 32H
-	+	θεριστής	-	θερίζω	+	+	+	-	+	-	-	-	-	θερίζομαι	+	-	faucher	32R2

- Exemples : fichier à part

Type de modifications



- modifications typographiques (55%)
 - ajout / suppression des espaces
 - modifications des caractères grecs/latins/majuscules/minuscules

N1VsupLocV-n -> N1 Vsup Loc V-n

- modifications au niveau de la structure des intitulés des propriétés (30%)
 - ajout / suppression des « =: », « , »

N1 = Npl obl -> N1 = : Npl obl

Type de modifications



- ajout des informations lexicales (9%)

- informations implicites $\emptyset \rightarrow VP$

- conventions différentes "*Pcomp*" $VNOσεDétN_2$
 - $P\acute{o}ti V NO σε N_2$
 - $Pνα V NO σε N_2$
 - *etc*

- suppression des colonnes (5%)

- non pertinentes/définitoires

$No Vsup N_1 Vadj \rightarrow \emptyset$

- conventions différentes

$Npréd \rightarrow Npred$

$Npréd \nearrow$

- changements linguistiques (1%)

Particularités du grec



- Nouveaux symboles

- Nouveau rôle thématique

- ✦ *Loc N2 =: (με+σε[me+se=avec]) N2 moyen-destination*

Αυτός κρύβει το ψωμί (με την+στην) πετσέτα

Il cache le pain avec la serviette

- Nouveaux traits sémantiques

- ✦ *N1 = : Nargent (υποτροφία [ypotrofía=bourse])*

Αυτός επένδυσε όλη την υποτροφία του σε ακίνητα

Il a investi toute sa bourse dans l'immobilier

Particularités du grec



- Mode des complétives

- plusieurs conjonctions demandent l'indicatif (*ὅτι* [*óti=que*], *πως* [*pos=que*], *που* [*rou=que*])
- une seule conjonction demande le subjonctif (*να* [*na=que*])

*Παραβλέπω **ὅτι** με κοροϊδεύεις (indicatif)*
** Παραβλέπω **πως** με κοροϊδεύεις (indicatif)*
*Je ne tiens pas compte **que** tu te moques de moi*

Solution : au lieu d'utiliser la notation *Pind* ou *Psubj* on utilise les notations *Pὅτι* [*óti=que*], *Pνα* [*na=que*] etc.

Particularités du grec

Problèmes des cas

ο Quelques règles générales :

- ✦ No : nominatif

Ο Πέτρος αμφισβητεί την αγάπη της

Pierre conteste son amour

No (nominatif) V N1

- ✦ N1 : accusatif

Εγώ ανοίγω την πόρτα

J'ouvre la porte

No V N1 (accusatif)

ο Exceptions des règles :

- ✦ No : accusatif, quand introduit par une préposition

Η αγάπη της αμφισβητείται απ' τον Πέτρο

Son amour est contesté par Pierre

N1 VP [από=par] No (accusatif)

- ✦ N1 : nominatif, dans le cadre de quelques transformations

Η πόρτα ανοίγει

La porte s'ouvre

N1 (nominatif) V

Particularités du grec

Problèmes des cas



○ D'autres problèmes

- ✦ *Κατά [katá] + groupe nominal au génitif*

Απάντησε κατά του πρωθυπουργού

Il s'est opposé au premier ministre

No V κατά [katá=contre] N2hum (génitif)

- ✦ *Κατά [katá] + groupe nominal à l'accusatif*

Μείωσε την κατανάλωση του γάλακτος κατά 250 γραμμάρια ημερησίως

Il a diminué sa consommation de lait de 250 grammes par jour

No V N1 κατά [katá=de] Dnum Nmes (accusatif)

Particularités du grec

Problèmes des cas



- Solution proposée :

3 règles applicables en priorité

1. tous les groupes nominaux introduits par une préposition sont à l'accusatif
2. le sujet (No) se trouve toujours au nominatif
3. le complément d'objet direct (N1) se trouve toujours à l'accusatif

το [to=le] Ρότι [ότι=que] VP (E+από [από=par] No)

Το ότι θα μειωθούν οι μισθοί αποφασίστηκε από το υπουργείο

Le fait que les salaires vont diminuer est décidé par le ministère

(No précédé par une préposition se trouve à l'accusatif)

No V

Ο Γιώργος ανησύχησε

Georges s'est inquiété (No est par défaut un nominatif)

Dans le cas contraire, le cas est précisé dans les tables du Lexique-Grammaire

No V κατά [katá = contre] N2hum:G

Απάντησε κατά του πρωθυπουργού

Il s'est opposé au premier ministre

No και [ke=et] N1:N VP

Ο Γιάννης και η Μαρία αγαπιούνται

Yannis et Maria s'aiment

Particularités du grec

Problèmes des cas



- verbes copules + attributs au nominatif
- verbes transitifs + complément à l'accusatif

N1hum:N είμαι [ime=être] Vpp

To κείμενο είναι γραμμένο (Le texte est écrit)

N1:N έχω [ého=avoir] V-n

Ο Γιάννης έχει στεναχώρια (Yannis a de la tristesse)

- Solution proposée :
 - dans un fichier à part nous citons tout verbe contenu dans l'ensemble des tables ainsi que le cas du groupe nominal qui suit en tant que complément ou en tant qu'attribut

είμαι [ime=être] : attribut au nominatif

βγάζω [vgázo=sortir]: complément à l'accusatif

Particularités du grec

Préfixes



- *Ξανα* [*ksaná=re*]-V

création d'une nouvelle entrée verbale en ajoutant le préfixe «*Ξανα*» en gardant toutes les propriétés et transformations du verbe sans le préfixe

Βγάζει το γάλα από το ψυγείο (Il sort le lait du frigo)

Ξαναβγάζει το γάλα από το ψυγείο (Il ressort le lait du frigo)

- NO *απο-* [*απο-=dés-*]V N1 Loc N2 source

ajout du préfixe "απο" (sans tiret) à l'entrée verbale (<ENT>V) pour former cette structure

εγκαθιστώ [*egkathistó=installer*]

Απεγκατέστησαν τους πυραύλους από το νησί

On a désinstallé les missiles de l'île

Table des classes



Propriété	table	V_38GL	V_32GCL
<ENT>V		o	o
Loc N lieu du passage		o	?
Loc N2 =: Loc N2 source, Loc # από		o	?
Loc N2 =: από N2 source		o	?
Loc N3 =: γύρω από N3 destination		o	?
Loc N3 =: πάνω σε N3 destination		o	?
Loc N3 =: σε N3 destination		o	?
N0 =: N-hum		?	o
N0 =: Nhum		+	o
N0 V N1		o	+
N0 V N1 (E+Loc N2 source) Loc N3 destination		o	?
N0 V N1 Loc N2 source (E+Loc N3 destination)		o	?
N0 V N1 Loc N2 source Loc N3 destination		+	?
N0 destination		o	?
N1 = Ppv =: (με+μας+σε+σας+τον+τους+τη+την+τις+το+τα)		o	+
N2 =: Nconc		+	?
ξανα-V		o	?
παρα-V		o	?

+ : propriété définitoire ou toujours valable pour les entrées de la table

- : propriété jamais valable pour les entrées de la table

o : propriété étudiée dans la table

? : propriété pas étudiée dans la table, à vérifier si valable

Documentation pour les auteurs

1	Present name of the feature	Value type	Former name of the feature 1 (if changed)	Former name of the feature 2 (if changed)	Classes where the feature is encoded	Linguistic comments	Information type	Argument concerned	Greek particularities in comparison to french
2	<ENT>V	lexical	Table6	Table9	30GEM, 30GEL, 32GCL, 32GRA, 6, 32GA, 9, 32GCV, 38GLH, 32GNM,		lemma	-	-
3	Loc N lieu du passage	binary	LocN=:lieudepassage	-	38GL	we don't know if it's an N2 or N3	??????	???	NOT USABLE FOR THE MOMENT
4	Loc N trajet	binary	LocNtrajet	-	38GL	we don't know if it's an N2 or N3	??????	????	NOT USABLE FOR THE MOMENT
5	Loc N2 = Ppv =:	binary	LocN2=ppv	Loc N2 = ppv	38GLD, 38GLH, 38GL,		relative	-	-
6	Loc N2 =: (με+σε) N2 moyen-destination	binary	Prép=:με/σε(moyen-dest)	Prép =: με / σε (moyen-dest)	38GLD, 38GL, 38GLR, 38GLS	moyen-destination: moyen AND destination AT THE	locative prepositions distribution+locative	N2	prepositions: με+σε
7	Loc N2 =: Loc N2 source, Loc # από	binary	Autresprépsrc	Autres prép src	38GLH, 38GL, 38GLS		locative prepositions distribution+locative	N2	prepositions: από
8	Loc N2 =: από N2 source	binary	Prép=:από(src)	Prép =: από (src)	38GLH, 38GL, 38GLS		locative prepositions distribution+locative	N2	prepositions: από
9	Loc N2 =: προς N2	binary	Prép=:προς(dir)	Prép =: προς (dir)	38GLD, 38GLH, 38GL, 38GLR, 38GLS		locative prepositions distribution+locative	N2	prepositions: προς
10	Loc N3 =: (από+σε) N3 destination	binary	Prép=:από/σε(dest)	Prép =: από / σε (dest)	38GLD, 38GLH, 38GL, 38GLR, 38GLS		distribution+locative prepositions	N3	prepositions: από+σε
11	Loc N3 =: Loc N3 destination, Loc # (σε+πάνω)	binary	Autresprépdest	Autres prép dest	38GLD, 38GLH, 38GL, 38GLR		locative prepositions distribution+locative	N3	prepositions: σε+πάνω σε+μέσα σε+γύρω
12	Loc N3 =: γύρω από N3 destination	binary	Prép=:γύρωαπό(dest)	Prép =: γύρω από (dest)	38GLD, 38GLH, 38GL, 38GLR		locative prepositions distribution+locative	N3	prepositions: γύρω από
13	Loc N3 =: μέσα σε N3 destination	binary	Prép=:μέσασε(dest)	Prép =: μέσα σε (dest)	38GLD, 38GLH, 38GL, 38GLR		locative prepositions distribution+locative	N3	prepositions: μέσα σε
14	Loc N3 =: μέχρι N3 destination	binary	Prép=:μέχρι(dest)	Prép =: μέχρι (dest)	38GLD, 38GLH, 38GL, 38GLR		locative prepositions distribution+locative	N3	prepositions: μέχρι
15	Loc N3 =: πάνω σε N3 destination	binary	Prép=:πάνωσε(dest)	Prép =: πάνω σε (dest)	38GLD, 38GLH, 38GL, 38GLR		locative prepositions distribution+locative	N3	prepositions: πάνω σε
16	Loc N3 =: σε N3 destination	binary	Prép=:σε(dest)	Prép =: σε (dest)	38GLD, 38GLH, 38GL, 38GLR		locative prepositions distribution+locative	N3	prepositions: σε
17	NO = PcompO =: το Pva	binary	NO=:το Pva	-	4		arguments distribution	N0	conjunctions: να

LGLex



Table des classes des verbes → possibilité d'extraire un **lexique syntaxique** des verbes à partir des tables (Constant & Tolone, 2010)

→ conversion depuis les tables Excel ou CSV grâce à l'outil *LGExtract*

→ format d'échange, reposant sur les mêmes concepts linguistiques que ceux manipulés dans les tables

→ format textuel ou XML

→ 3 966 entrées verbales, 4 252 en comptant les verbes préfixés

LGLex



- **ID=catégorie_numTable_numEntrée** : identifiant indiquant sa catégorie, la table dont il provient et le numéro de l'entrée dans cette table
- **lexical-info** : informations lexicales liées à l'entrée (lemme et prépositions associées à certains arguments)
- **args** : distributions des arguments, avec éventuellement d'autres informations (traits sémantiques, mode et contrôle de la complétive, prépositions)
- **all-constructions** : constructions possibles (soit nommées de façon complète avec tous les éléments dans l'ordre, soit des transformations à partir de construction de base)

LGLex



ID=V_38GL_33

lexical-info=[**cat**="verb",verb=[lemma="βγάζω"],pfx-V=(verb="ξαναβγάζω",verb="παραβγάζω"),
prepositions=(),locatifs=(locatif=[id="2",list=(prep="από")],
locatif=[id="3",list=(prep="σε")])]

args=(const=[**pos**="ο",dist=(comp=[**cat**="NP",**hum**="true",introd-prep=(),introd-loc=(),
origin=(orig="No =: Nhum")])],

const=[**pos**="1",dist=(comp=[**cat**="NP",**conc**="true",introd-prep=(),introd-loc=(),
origin=(orig="N1 =: Nconc")])])]

const=[**pos**="2",dist=(comp=[**cat**="NP",**conc**="true",introd-prep=(),introd-loc=(),
origin=(orig="N2 =: Nconc")])],

all-constructions=[absolute=(construction="true::No V N1 Loc N2 source Loc N3 destination",
construction="o::No V N1 Loc N2 source (E+Loc N3
destination)",

construction="o::No V N1 (E+Loc N2 source) Loc N3

destination"),

relative=(construction="N1 = Ppv =:

(με+μας+σε+σας+τον+τους+τη+την+τις+το+τα)")]

Perspectives



- **Tables des classes :**
 - Notation des propriétés qui sont toujours ou jamais valables pour l'ensemble des entrées d'une table
- **Tables**
 - Exemples
 - Traduction → études bilingues
 - Nouvelles tables → conventions de notation du grec

Perspectives



- Conversion du lexique *LGLex* au format *Lefff* (Sagot, 2010), comme cela a été fait pour le français (Tolone & Sagot, 2011)
- Adaptation de la méta-grammaire FRMG (Thomasset & de la Clergerie, 2005) du français à celle du grec, comme cela a été fait pour l'espagnol (Fernandez, 2010)
- Application de la même démarche aux noms prédicatifs (de la langue générale et de la langue terminologique) (Sfetsiou, 2007)

Références



- Constant, M. & Tolone, E. (2010). A generic tool to generate a lexicon for NLP from Lexicon-Grammar tables. In Michele De Gioia, editor, *Actes du 27e Colloque international sur le lexique et la grammaire (L'Aquila, 10-13 septembre 2008). Seconde partie*. Volume 1 of *Lingue d'Europa e del Mediterraneo, Grammatica comparata*, pages 79-93. Aracne.
- Fernández González, D. (2010). *Cadena de procesamiento lingüístico para el español*. Mémoire de D.E.A., Université de Vigo, Espagne.
- Fista, E. (2007). *Μέθοδος αυτόματης αναπαράστασης προθηματοποιημένων και σύνθετων ρημάτων της Νέας Ελληνικής*. Thèse de doctorat, Université Aristote de Thessalonique.
- Gross, M. (1975). *Méthodes en syntaxe*. Paris: Hermann.
- Ioannidou, K. & Kyriacopoulou, T. (2010). Lexique-grammaire des verbes en grec moderne : Ambiguités formelles et analyse des textes. *Actes du 29ème Colloque International sur le Lexique et la Grammaire*, pp. 151-160. Belgrade, Serbie : Faculty of Mathematics of University of Belgrade.
- Kyriacopoulou, T. (2005). *L'analyse automatique des textes écrits : le cas du grec moderne*. Thessalonique: University Studio Press.
- Sagot, B. (2010). *The Lefff, a freely available and large-coverage morphological and syntactic lexicon for French*. Proceedings of the 7th Language Resources and Evaluation Conference, La Valette, Malte.
- Sfetsiou, V. (2007). *Κατηγορηματικά ονόματα: μέθοδος ανάλυσής τους για ηλεκτρονικές εφαρμογές*. Thèse de doctorat, Université Aristote de Thessalonique.
- Thomasset, F & de La Clergerie, É. (2005). Comment obtenir plus des Méta-Grammaires, *Actes de la Conférence sur le Traitement Automatique des Langues Naturelles*. Dourdan, France.
- Tolone, E. (2011). *Analyse syntaxique à l'aide des tables du Lexique-Grammaire du français*. Thèse de doctorat, Université Paris-Est, 326 pp.
- Tolone, E. & Sagot, B. (2011). Using Lexicon-Grammar tables for French verbs in a large-coverage parser. In Z. Vetulani, editor, *Human Language Technology, Forth Language and Technology Conference, LTC 2009, Poznań, Poland, November 2009, Revised Selected Papers. Lecture Notes in Artificial Intelligence*. Springer Verlag. À paraître.
- Voskaki, O. (2011). *Le lexique-grammaire des verbes du grec moderne : Constructions transitives non locatives à un complément d'objet direct*. Thèse de doctorat, Université Paris-Est.
- Yannacopoulou, A. (2005). *Le lexique-grammaire des verbes du grec moderne : les constructions transitives locatives standard*. Thèse de doctorat, Université de Marne-la-Vallée.