

Acquiring Typed Predicate-Argument Structures from Corpora

Elisabetta Jezeq
Università di Pavia

International Workshop
WordNet and Sketch Engine
Pavia, March 9, 2012

- The goal of the presentation is to illustrate the methodology we are currently using to acquire typed predicate-argument structures from corpora, with the aim of compiling a repository of corpus-based patterns for Italian verbs (PDIV).

- The goal of the presentation is to illustrate the methodology we are currently using to acquire typed predicate-argument structures from corpora, with the aim of compiling a repository of corpus-based patterns for Italian verbs (PDIV).
- This methodology involves using Sketch Engine, a corpus query system incorporating word sketches, one-page automatic summary of a word's grammatical and collocational behaviour.

- The goal of the presentation is to illustrate the methodology we are currently using to acquire typed predicate-argument structures from corpora, with the aim of compiling a repository of corpus-based patterns for Italian verbs (PDIV).
- This methodology involves using Sketch Engine, a corpus query system incorporating word sketches, one-page automatic summary of a word's grammatical and collocational behaviour.
- Examine the main problems that arise in the identification of the frames.

Repository of typed predicate-argument structures for Italian 1/2

- Lexical resource under development with the technical support of the Faculty of Informatics at Masaryk University in Brno (CZ).

Repository of typed predicate-argument structures for Italian 1/2

- Lexical resource under development with the technical support of the Faculty of Informatics at Masaryk University in Brno (CZ).
- It currently consists of a small nucleus of 50 lexical units (verbs).

Repository of typed predicate-argument structures for Italian 1/2

- Lexical resource under development with the technical support of the Faculty of Informatics at Masaryk University in Brno (CZ).
- It currently consists of a small nucleus of 50 lexical units (verbs).
- It is a manually annotated resource.

Repository of typed predicate-argument structures for Italian 1/2

- Lexical resource under development with the technical support of the Faculty of Informatics at Masaryk University in Brno (CZ).
- It currently consists of a small nucleus of 50 lexical units (verbs).
- It is a manually annotated resource.
- Linguistic research and computational applications.

Repository of typed predicate-argument structures for Italian 2/2

- In this resource, each lexical unit (V) is linked to a set of patterns/frames extracted from the corpus following the Corpus Pattern Analysis methodology (CPA, Hanks - Pustejovsky 2005).

Repository of typed predicate-argument structures for Italian 2/2

- In this resource, each lexical unit (V) is linked to a set of patterns/frames extracted from the corpus following the Corpus Pattern Analysis methodology (CPA, Hanks - Pustejovsky 2005).
- Each frame is associated with a verb sense (occasionally, one sense is associated with multiple frames).

Repository of typed predicate-argument structures for Italian 2/2

- In this resource, each lexical unit (V) is linked to a set of patterns/frames extracted from the corpus following the Corpus Pattern Analysis methodology (CPA, Hanks - Pustejovsky 2005).
- Each frame is associated with a verb sense (occasionally, one sense is associated with multiple frames).
- Each frame is associated with a set of annotated corpus instances (a sample of 250 occurrences for each verb), that represent more prototypical and less prototypical instantiations of the frame.

Repository of typed predicate-argument structures for Italian 2/2

- In this resource, each lexical unit (V) is linked to a set of patterns/frames extracted from the corpus following the Corpus Pattern Analysis methodology (CPA, Hanks - Pustejovsky 2005).
- Each frame is associated with a verb sense (occasionally, one sense is associated with multiple frames).
- Each frame is associated with a set of annotated corpus instances (a sample of 250 occurrences for each verb), that represent more prototypical and less prototypical instantiations of the frame.
- <http://nlp.fi.muni.cz/projects/cpa/>

What counts as a typed predicate-argument structure?

A typed predicate-argument structure is a corpus-derived verb frame with the specification of the expected semantic type for each argument position, populated by lexical sets, i.e. the statistically relevant list of collocates that typically fill each position.

[[Persona]] partecipa a [[Evento]]

- Lexical set [[Evento]] = {gara, riunione, selezione, manifestazione, seduta, cerimonia, conferenza, votazione, elezione, celebrazione, esequia, competizione, maratona, discussione, messa, festa, marcia, fiaccolata, trattativa, missione, commemorazione, incontro, concorso, convegno, raduno, iniziativa, stage, evento, seminario, torneo, attività, corso, asta, dibattito, progetto, festival. . . }

Tools needed to acquire typed predicate-argument structures

- A balanced corpus of the language (i.e. general language).
Reduced version of ITWaC (Baroni & Kilgarriff 2006).

Tools needed to acquire typed predicate-argument structures

- A balanced corpus of the language (i.e. general language).
Reduced version of ITWaC (Baroni & Kilgarriff 2006).
- A methodology of corpus analysis.
CPA, Corpus Pattern Analysis, Hanks - Pustejovsky 2005.
Including statistical corpus analysis, Church - Hanks 1989.

Tools needed to acquire typed predicate-argument structures

- A balanced corpus of the language (i.e. general language).
[Reduced version of ITWaC \(Baroni & Kilgarriff 2006\).](#)
- A methodology of corpus analysis.
[CPA, Corpus Pattern Analysis, Hanks - Pustejovsky 2005.](#)
[Including statistical corpus analysis, Church - Hanks 1989.](#)
- A model of compositional mechanisms that are active in V-ARG selection (type coercion etc.).
[Pustejovsky 2006.](#)

Tools needed to acquire typed predicate-argument structures

- A balanced corpus of the language (i.e. general language).
[Reduced version of ITWaC \(Baroni & Kilgarriff 2006\).](#)
- A methodology of corpus analysis.
[CPA, Corpus Pattern Analysis, Hanks - Pustejovsky 2005.](#)
[Including statistical corpus analysis, Church - Hanks 1989.](#)
- A model of compositional mechanisms that are active in V-ARG selection (type coercion etc.).
[Pustejovsky 2006.](#)
- A “shallow” list of semantic type labels (HUMAN, ARTEFACT, EVENT, ecc.) borrowed from the English project.
<http://deb.fi.muni.cz/pdev/>

Tools needed to acquire typed predicate-argument structures

- A balanced corpus of the language (i.e. general language).
[Reduced version of ITWaC \(Baroni & Kilgarriff 2006\).](#)
- A methodology of corpus analysis.
[CPA, Corpus Pattern Analysis, Hanks - Pustejovsky 2005.](#)
[Including statistical corpus analysis, Church - Hanks 1989.](#)
- A model of compositional mechanisms that are active in V-ARG selection (type coercion etc.).
[Pustejovsky 2006.](#)
- A “shallow” list of semantic type labels (HUMAN, ARTEFACT, EVENT, ecc.) borrowed from the English project.
<http://deb.fi.muni.cz/pdev/>
- A suite of corpus tools: Manatee, Bonito, Sketch Engine.
[Kilgarriff, Rychly, Smrz, Tugwell 2004.](#)

Procedure

- Create a sample concordance (KWIC index) for the target verb using SkE.
 - 250 examples of actual uses of the target verb.

- Create a sample concordance (KWIC index) for the target verb using SkE.
 - 250 examples of actual uses of the target verb.
- Analyze the lexical set for each grammatical relation (SUBJ_OF, OBJ_OF, IOBJ_OF ...) through the word sketches.

- Create a sample concordance (KWIC index) for the target verb using SkE.
 - 250 examples of actual uses of the target verb.
- Analyze the lexical set for each grammatical relation (SUBJ_OF, OBJ_OF, IOBJ_OF ...) through the word sketches.
- Manual clustering of lexical sets.

- Create a sample concordance (KWIC index) for the target verb using SkE.
 - 250 examples of actual uses of the target verb.
- Analyze the lexical set for each grammatical relation (SUBJ_OF, OBJ_OF, IOBJ_OF ...) through the word sketches.
- Manual clustering of lexical sets.
- Identify the frames (typical syntagmatic patterns).

- Create a sample concordance (KWIC index) for the target verb using SkE.
 - 250 examples of actual uses of the target verb.
- Analyze the lexical set for each grammatical relation (SUBJ_OF, OBJ_OF, IOBJ_OF ...) through the word sketches.
- Manual clustering of lexical sets.
- Identify the frames (typical syntagmatic patterns).
- Assign each instance of the sample to one of the frames.

- Create a sample concordance (KWIC index) for the target verb using SkE.
 - 250 examples of actual uses of the target verb.
- Analyze the lexical set for each grammatical relation (SUBJ_OF, OBJ_OF, IOBJ_OF ...) through the word sketches.
- Manual clustering of lexical sets.
- Identify the frames (typical syntagmatic patterns).
- Assign each instance of the sample to one of the frames.
- Store the frames in the entry manager.

- Create a sample concordance (KWIC index) for the target verb using SkE.
 - 250 examples of actual uses of the target verb.
- Analyze the lexical set for each grammatical relation (SUBJ_OF, OBJ_OF, IOBJ_OF ...) through the word sketches.
- Manual clustering of lexical sets.
- Identify the frames (typical syntagmatic patterns).
- Assign each instance of the sample to one of the frames.
- Store the frames in the entry manager.
- Associate each frame with one verb sense; the sense is “anchored” to the frame through the semantic types.

- Create a sample concordance (KWIC index) for the target verb using SkE.
 - 250 examples of actual uses of the target verb.
- Analyze the lexical set for each grammatical relation (SUBJ_OF, OBJ_OF, IOBJ_OF ...) through the word sketches.
- Manual clustering of lexical sets.
- Identify the frames (typical syntagmatic patterns).
- Assign each instance of the sample to one of the frames.
- Store the frames in the entry manager.
- Associate each frame with one verb sense; the sense is “anchored” to the frame through the semantic types.
- [[Persona]] prendere parte a [[Evento]].

Unclassified sample from ITWaC for the verb *divorare*

- Gianfranco , ti aggiorno che oltre ai dolci ... **divoro** anche le verdure.
- la condanna viene eseguita, un incendio **divora** Mahagonny.
- si incolla, il gelato squagliato, cibi **divorati** distrattamente, senza un pensiero.
- quinta tappa, da Blois a Montargis, **è stata divorata** alla incredibile media di 48,584.
- capace di acchiappare al volo i neofiti (che **divorati** dalla voglia di leggere i numeri passati...
- rimanendo di guardia affinché la grafica non **divori** la fantasia artistica
- Una volta uccisi i nemici, tendono a **divorarli** per acquisirne il potere.
- invece direttamente nella polpa, raggiunge e **divora** i semi. Gli attacchi del fitofago su melo
- posso approvare che la bestia più grande **divori** la bestia più piccola
- coinvolgere dalla storia e l' **ho** letteralmente **divorato**. Questo libro dura quasi 50 anni

divorare

- [[Animate]-subj] divora [[Food]-obj]

divorare

- [[Animate]-subj] divora [[Food]-obj]
- [[Human]-subj] divora [[Document]-obj]

divorare

- [[Animate]-subj] divora [[Food]-obj]
- [[Human]-subj] divora [[Document]-obj]
- [[Vehicle]-subj] divora [[Path]-obj]

divorare

- [[Animate]-subj] divora [[Food]-obj]
- [[Human]-subj] divora [[Document]-obj]
- [[Vehicle]-subj] divora [[Path]-obj]
- [[Emotion]-subj] divora [[Human]-obj]

divorare

- [[Animate]-subj] divora [[Food]-obj]
- [[Human]-subj] divora [[Document]-obj]
- [[Vehicle]-subj] divora [[Path]-obj]
- [[Emotion]-subj] divora [[Human]-obj]
-

Patterns for the verb *divorare*

The screenshot shows a window titled "divorare: CPA Patterns". The window has a green header bar with the text "Patterns for: divorare" and several buttons: "Add", "Copy", "Corpora", "Preview", "Renumber", "Delete", and "Close". Below the header bar, there is a "Save" button, two input fields for "Sample size" and "Semantic class", and the text "Erlangen No". The main area of the window contains a list of five patterns, each with a checkbox and a number:

- 1 **[[Animate]] divorare [[Food]]**
[[Animate]] ingest and consume [[Food]] quickly and eagerly
- 2 **[[Human]] divorare [[Document]]**
[[Human]] read [[Document]] quickly and eagerly
- 3 **[[Vehicle]] divorare [[Path]]**
[[Vehicle]] cover [[Path]] quickly
- 4 **[[Emotion]] divorare [[Human]]**
[[Emotion]] consume, absorb [[Human]]
- 5 **[[Event]] divorare [[Physical Object]]**
[[Event]] destroy [[Physical Object]] completely

Pattern1 [[Animate]-subj] divorà [[Food]-obj]

- quando nessuno lo vedeva, **divorò** il salame da solo
- Mentre **divoravano** la cena
- **l'aquila** che **divora** il **serpente**
- Cinzia non **divorò** il panettone
- Durante la festa i cittadini **hanno divorato** in pochi secondi una torta di 8
- E **le cavallette**, buone solo a **divorare** i raccolti?
- È molto magra, **divora** i biscotti con gran voracità
- Nascostamente **le larve avevano divorato i bordi delle tele**
- davvero molto bellina!! Oltre a **divorare** gelati (d'estate me ne mangio
- Bevono caffelatte, **divorano** cornetti, comprano
- **Il leone** che assalta e **divora l'uomo** è un'offesa per la ragione
- sorseggiare un caffè in tranquillità o **divorare** un panino
- **bruchi**, che minacciano di **divorare le foglie** delle vostre piante
- a guardare la televisione a bere birra e **divorare** patatine.

Pattern2: [[Human]-subj] divora [[Document]-obj]

- di mettersi a studiare, a **divorare** libri come non mai.
- e appena ho avuto la possibilità **ho divorato** il libro.
- L'amica **divorò** il romanzo in meno di tre giorni.
- Durante le medie **ho divorato** i racconti della "Golden Age".
- Nelle ore libere **divoravo** i grossi e piccoli volumi di storia.
- Legge soprattutto romanzi e **divora** i fumetti, sua vera passione.
- Spinti da genitori e insegnanti a **divorare** intere enciclopedie
- **un autore** da **divorare** a cui affidare il nostro recondito.
- ma soprattutto **ho** scoperto e **divorato la Deledda**, mi piace da impazzire.

Pattern3: [[Vehicle]-subj] **divora** [[Path]-obj]

- sono migliorati a ogni uscita, **hanno divorato** i chilometri.
- **Ha divorato** la pista, la conosceva bene.
- la Maserati **divora** i rettilinei di Adria.
- Il poderoso incrociatore che **divorava** miglia su miglia.
- Il gommone **divora** 71 chilometri sulle onde e lì
- in certe gare **il vincitore divora** 90 chilometri in sette ore.
- la Skoda **divora** chilometri.

Pattern4: [[Emotion]-subj] divorava [[Human]-obj]

- frenare la gelosia che ormai lo **divorava**.
- È meglio così. Il tuo amore ti **avrebbe divorato**.
- altrimenti l'ansia e l'angoscia vi **divorerà**.
- Un atroce dubbio mi **divora** e nessuno riesce a scioglierlo.
- sotto lo stimolo della cupidigia che lo **divorava**, sarebbe sicuramente arrivato
- il senso di colpa **divora** il boia Kaliyappan
- con i sensi di colpa che ci **divorano**.
- a combattere la depressione che la **divorava** da anni.
- La tensione mi **sta divorando**.

Contexts Pattern1 *divorare*

Annotating: divorare [Info](#) [Sort](#) [Finish](#) New pattern: Number globally:

#76962398	piatto di pasta con i funghi , mi guardavi e	divoravi	il	quel piatto , stavo andando via quando
#94270451	scampare al mare , il pericolo di finire	divorati	il	dai selvaggi sarebbe stato superiore alla
#121260390	passione è mangiare ; sei come una termite :	divori	il	tutto il commestibile che ti capita a portata
#125028431	attorno come leone ruggente , cercando di	divorare	il	, ma la sua attività è frenata durante
#126504501	mangiare , poi pokerino , poi in 6 s' è	divorato	il	9 chili di cocomero (anche se nn mi piace
#126517710	fantasia e nell' orrore nel vedere una persona	divorare	il	i resti di un animale con piacere e indemoniata
#130157369	vita della crosta terrestre . I batteri che	divorano	il	il petrolio rendono inutilizzabili la maggior
#140386566	. S' accampano ai piedi di alcuni pini	divorando	il	radici e licheni per ingannare la fame
#163932620	per raggiungerle , così poi da ucciderli e	divorarli	il	. Nell' Odissea le Sirene sono legate a
#173602762	dalle carni significano allora cessare di	divorare	il	la carne del povero , lottare perché tutti
#176746065	Qualche prete , vedendolo sbrindellato e	divorato	il	dagli insetti , non ardiva dargli l' ostia
#177788496	. La sera compriamo delle provviste che	divoriamo	il	sul nostro terrazzino . Ed eccoci alla
#179514820	posture , gesti grotteschi dei cadaveri	divorati	il	sotto il sole da schiere di mosche " .
#201358385	sufficienti a sfamare la ragazza , ma lei le	divorò	il	in un baleno . " Le hai dato tutte le nostre
#219191349	delle mani , nessuno del cucchiaino . Chi	divorava	il	più in fretta riceveva di più ; il più
#219581855	guardano mentre quello preso in pescheria lo	divorano	il	!!! In bocca al lupo DeIorenziDanilo Inserito
#239376048	lei siamo degli spietati assassini , che	divorano	il	anatre , conigli , piccioni , rane , pesci
#281775601	... Guarda quel brutto giaguaro che sta	divorando	il	la testa della tua vittima . " Prutta
#281775803	erano gettate anche sul giaguaro , e lo	divoravano	il	ancora agonizzante . Piene fino a scoppiare
#339203416	reticella che impedisca ai riproduttori di	divorare	il	le uova ; ci sono da tenere presente 2
#341664477	noi si sono estese a vista d' occhio : ora	divoravano	il	l' erba su di un circolo irregolare di

- The paradigmatic sets of words that populate specific argument slots within the same verb sense do not map neatly onto the “expected” type (selected by V).

- The paradigmatic sets of words that populate specific argument slots within the same verb sense do not map neatly onto the “expected” type (selected by V).
- Mismatch between “target” type (assigned by V) and “source” type (of N) within the same grammatical relation.

- The paradigmatic sets of words that populate specific argument slots within the same verb sense do not map neatly onto the “expected” type (selected by V).
- Mismatch between “target” type (assigned by V) and “source” type (of N) within the same grammatical relation.
- Regular alternation of semantic type within the same argument position.

Type mismatches

- [[Vehicle]–subj] atterra
L'aereo | il pilota | il Papa | il volo stava atterrando.

Type mismatches

- [[Vehicle]–subj] atterra
L'aereo | **il pilota** | **il Papa** | **il volo** stava atterrando.
- [[Human]–subj] commenta, pensa, crede, annuncia ...
Il Presidente | **il Quirinale** | **il governo** non ha commentato.
Lo speaker | **l'altoparlante** | **il campanello** annunciava l'arrivo del treno.
I negozianti | **I cartelli colorati** annunciano sconti.

Type mismatches

- [[Vehicle]–subj] atterra
L'aereo | **il pilota** | **il Papa** | **il volo** stava atterrando.
- [[Human]–subj] commenta, pensa, crede, annuncia ...
Il Presidente | **il Quirinale** | **il governo** non ha commentato.
Lo speaker | **l'altoparlante** | **il campanello** annunciava l'arrivo del treno.
I negozianti | **I cartelli colorati** annunciano sconti.
- [[Human]–subj] sanguina
L'uomo | **la mia mano** ha smesso di sanguinare.

Type mismatches

- [[Vehicle]–subj] atterra
L'aereo | **il pilota** | **il Papa** | **il volo** stava atterrando.
- [[Human]–subj] commenta, pensa, crede, annuncia ...
Il Presidente | **il Quirinale** | **il governo** non ha commentato.
Lo speaker | **l'altoparlante** | **il campanello** annunciava l'arrivo del treno.
I negozianti | **I cartelli colorati** annunciano sconti.
- [[Human]–subj] sanguina
L'uomo | **la mia mano** ha smesso di sanguinare.
- [[Human]–subj] divora [[Documento]-obj]
divorare un romanzo | **un autore** | **Freud**.

- “Regular choices of types within an overall pattern are coded as type alternations. Common alternations in subject position are for instance `[[Human|Institution]]` and `[[Human|Body Part]]`”.

- “Regular choices of types within an overall pattern are coded as type alternations. Common alternations in subject position are for instance `[[Human|Institution]]` and `[[Human|Body Part]]`”.

- “Regular choices of types within an overall pattern are coded as type alternations. Common alternations in subject position are for instance `[[Human|Institution]]` and `[[Human|Body Part]]`”.
- “Non-canonical lexical items breaking a particular statistical threshold are “coerced” into ‘honorary’ membership of a semantic type in particular contexts.”

Some of the words that are potential candidates for the lexical set of the “target” type are not instantiated.

[[Persona]] partecipa a [[Evento]]

- Lexical set [[Evento]] = {gara, riunione, selezione, manifestazione, seduta, cerimonia, conferenza, votazione, elezione, celebrazione, esequia, competizione, maratona, discussione, messa, festa, marcia, fiaccolata, trattativa, missione, commemorazione, incontro, concorso, convegno, raduno, iniziativa, stage, evento, seminario, torneo, attività, corso, asta, dibattito, progetto, festival... }
- ... but not [[Evento]] = {temporale, suicidio... }

- Lexical sets populating a node in the ontology (i.e. a semantic type) tend to “shimmer” – that is, the membership of the lexical set tends to vary when one moves from verb to verb: some words drop out while other come in, just as predicated by Wittgenstein (*family resemblances*).

Shimmering Lexical Sets

- Lexical sets populating a node in the ontology (i.e. a semantic type) tend to “shimmer” – that is, the membership of the lexical set tends to vary when one moves from verb to verb: some words drop out while other come in, just as predicated by Wittgenstein (*family resemblances*).
- Different verbs select different prototypical members of a semantic type even if the rest of the set remains the same.

lavare [[Body Part]-obj]

- Lexical set **[[Body Part]]** = {denti, mano, piede, viso, faccia, schiena, testa, orecchio, volto ... }

amputare [[Body Part]-obj]

- Lexical set **[[Body Part]]** = {arto, gamba, braccio, dito, orecchio, mano, piede ... }

What is a [[Document]]?

leggere

{opinione, commento, libro, avvertenza, recensione, giornale, cronaca, intervista, articolo, blog, messaggio, poesia, notizia, resoconto, racconto, fumetto, romanzo ... }

pubblicare

{articolo, intervista, nota, avviso, bando, studio, libro, sondaggio, notizia, volume, saggio, testo, giornale, comunicato, monografia, rivista, documento, inesattezza, annuncio ... }

spedire

{cartolina, pacco, e-mail, fax, messaggio, lettera, raccomandata, foto, telegramma, copia, sms, merce curriculum, pacchetto, invito, vaglia, posta, file, libro, coupon, modulo ... }

tradurre

{testo, frase, bibbia, parola, brano, libro, poesia, vocabolo, concetto, opera, versetto, termine, idea, nome, spiegazione, romanzo, canzone, pensiero, vangelo, espressione ... }

- “The ontology is not conceived as a rigid yes/no structure, but as a statistically based structure of collocational preferences, which we call “shimmering lexical sets” .
- Hanks - Jezek 2008.

- ... book <read 772/9037; write 933/13015; publish 416/7230, borrow 43/1358 ... >

- ... book <read 772/9037; write 933/13015; publish 416/7230, borrow 43/1358 ... >
- ... novel <write 182/13015; read 88/9037; publish 45/7230; set 19/14528 ... >

- ... book <read 772/9037; write 933/13015; publish 416/7230, borrow 43/1358 ... >
- ... novel <write 182/13015; read 88/9037; publish 45/7230; set 19/14528 ... >
- ... article <write 263/13015; publish 174/7230; read 156/9037; contribute 28/1313 ... >

- ... book <read 772/9037; write 933/13015; publish 416/7230, borrow 43/1358 ... >
- ... novel <write 182/13015; read 88/9037; publish 45/7230; set 19/14528 ... >
- ... article <write 263/13015; publish 174/7230; read 156/9037; contribute 28/1313 ... >
- ... letter <write 1032/13015; send 540/12011; receive 544/18381; post 77/451 ... >

Complex Types

VERBS may select complex “target” types.

NOUNS may be associated with complex “source” types.

[[Human]-subj] legge [[Phys•Info]-obj]

- [[Phys•Info]]: {libro, lettera ...}
- [[Info]]: {intervista, discorso, musica ...}
- [[Phys]]: {contatore (?)}

[[Human]-subj] raggiunge [[Location]-obj]

- [[Location]]: {riva del mare, vetta, località turistica, radura ...}
- [[Artefact]]: {cancello, macchina, semaforo ...}
- [[Artefact•Location]]: {casa, albergo, aeroporto ...}

- Create a repository of corpus-based patterns intended as *semantically motivated syntagmatic distinctions*.

- Create a repository of corpus-based patterns intended as *semantically motivated syntagmatic distinctions*.
- Build an ontology for NOUNS, grouped into clusters (shimmering lexical sets) according to their argument roles and their ability to predict the meaning of V in context.

- Create a repository of corpus-based patterns intended as *semantically motivated syntagmatic distinctions*.
- Build an ontology for NOUNS, grouped into clusters (shimmering lexical sets) according to their argument roles and their ability to predict the meaning of V in context.
- Compile an inventory of metonymic shifts (metonymic map).

- Use typed predicate-argument structures acquired from corpora to enrich existing lexicographic resources and improve their performance in NLP tasks.

- Use typed predicate-argument structures acquired from corpora to enrich existing lexicographic resources and improve their performance in NLP tasks.
- For example, sense distinctions for Vs induced from empirical data through large-scale corpus investigation improve the chances to achieve broad coverage WSD.

- Develop a classification of Vs based on the argument typing (orthogonal to Levin's and VerbNet classes).

- Develop a classification of Vs based on the argument typing (orthogonal to Levin's and VerbNet classes).
- [[Human]-subj]_[[Location]-obj]: *raggiungere*, *attraversare*, *costeggiare* (motion verbs), but also *invadere*, *bombardare*, *conquistare*, *visitare* ...

- Develop a classification of Vs based on the argument typing (orthogonal to Levin's and VerbNet classes).
- [[Human]-subj]-[[Location]-obj]: *raggiungere*, *attraversare*, *costeggiare* (motion verbs), but also *invadere*, *bombardare*, *conquistare*, *visitare* ...
- [[Human]-subj]-[[Event]-obj]: *finire*, *iniziare*, *interrompere* (aspectual verbs), but also *organizzare*, *vietare*, *proibire* ...

- Develop a classification of Vs based on the argument typing (orthogonal to Levin's and VerbNet classes).
- [[Human]-subj]_[[Location]-obj]: *raggiungere, attraversare, costeggiare* (motion verbs), but also *invadere, bombardare, conquistare, visitare* ...
- [[Human]-subj]_[[Event]-obj]: *finire, iniziare, interrompere* (aspectual verbs), but also *organizzare, vietare, proibire* ...
- Class members are not necessarily synonyms or antonyms.

- Baroni, M. and A. Kilgarriff 2006. “Large Linguistically-Processed Web Corpora for Multiple Languages”, *EACL 2006 Proceedings*, 87-90.
- Church, K.W. and P. Hanks 1990. “Word Association Norms, Mutual Information and Lexicography”. In *Computational Linguistics*, vol. 16, n. 1, 22-29.
- Hanks, P. and E. Jezek 2008. “Shimmering Lexical Sets”. In E. Bernal and J. De Cesaris (eds) *Proceedings of the XIII EURALEX International Congress*, Barcelona, Universidad Pompeu Fabra, 391-402.
- Hanks, P. and J. Pustejovsky 2005. “A Pattern Dictionary for Natural Language Processing”. In *Revue française de linguistique appliquée*, 10 (2).

- Kilgarriff A., P. Rychly, P. Smrz & D. Tugwell 2004. “The Sketch Engine”. In G. Williams and S. Vessier (eds.). *Proceedings of the XI EURALEX International Congress* (July 6-10, 2004). Lorient. France. 105-11.
- Pustejovsky, J. 2006. “Type theory and lexical decomposition”. In *Journal of Cognitive Science* 6, 39-76.
- Pustejovsky, J. and E. Jezek 2008. “Semantic Coercion in Language: Beyond Distributional Analysis”. In *Italian Journal of Linguistics* 20:1, 181-214.