

Report for the STSM (Potsdam 2015)

Margot Colinet
ALPAGE-INRIA
Université Paris Diderot-Paris 7

June 30, 2015

1 DiMLexConn

The goal of this short term collaborative project between the French Discourse Treebank Project (*alpage*, Université Paris Diderot-Paris 7) and the Applied CompLing Discourse Research Lab (Universität Potsdam), was a comparative study of French and German lexicons of discourse connectives, French LexConn (Roze et al 2012) and the German DiMLex (Stede, 2002).

As a starting point, I created a (French-German) bilingual list of discourse connectives. The list contains every entries of LexConn which are translated (if possible) with a corresponding entry of DiMLex.

Example of an entry in the bilingual (French-German) list of discourse connectives (DimLexConn) (xml format):

```
<connective cat="subj" id="c37" relations="condition" type="sub">
  <forme>aussi longtemps que</forme>
  <correspondances>
    <deconn dimlexid="k267" cat="subj" RSTrel="" SDRTrel="">
      solange
      <examples>
        <frexample>Je reste *aussi longtemps qu'* il pleut.</frexample>
        <deexample>Ich bleibe *solange* es regnet.</deexample>
        <source>Europarl3</source>
      </examples>
    </deconn>
  </correspondances>
</connective>
```

First of all, I listed automatically first the numerous entries of the French lexicon LexConn that couldn't be match with any entry in DimLex. Then, I listed automatically all the connectives from German lexicon DiMLex that couldn't correspond to any entry in LexConn. From these data, I could detect a number of general mismatches between LexConn and DiMLex, that I will briefly present in the following document.

2 General mismatches

- A point specific to DiMLex: the prepositions (+noun).
- A point specific to German: the pronominal adverbs.
- A point specific to French: prepositions + infinitival clauses.
- A point specific to French: NP and PP as lexicalised adverbial connectives.

2.1 Prepositions + noun

The first difference lies on the syntactic definition of an discourse argument. In LexConn, only sentences are considered as arguments, while in Dimlex eventive nouns are also considered as arguments.

As a consequence, in LexConn, only occurrences of prepositions that can take an infinitival clause as an argument have been annotated. The (automatic) annotation consisted in eliminating the occurrences that took a NP as an argument (like (2), (3) and (4)), despite the fact that the preposition *jusqu'à* (*until*) in (2) might be a connective taking an eventive noun as its argument:

- (1) Paul a couru **jusqu'à** (en) être épuisé.
"Paul ran up to be exhausted."
- (2) Paul a couru **jusqu'à** épuisement.
"Paul ran up to exhaustion."
- (3) Paul a couru **jusqu'à** moi.
"Paul ran up to me."
- (4) Elle a rempli son verre **jusqu'au** bord.
"She filled her glass up to the top."

There are two kinds of prepositions in LexConn: (i) the ones which can take only an infinitival clause as an argument, like *avant de* (*before*), see (5), (ii) the three prepositions which can take either a noun or an infinitival clause as an argument, see (7) and (8) :

- *pour* (*in order to/for*)
 - *jusqu'à* (*until*)
 - *après* (*after*)
- (5) Il a pris une douche **avant de** sortir.
'He took a shower before going out.'
 - (6) Il s'est allongé **afin de** se reposer.
'He laid to rest.'
 - (7) Il a cuisiné **pour** faire plaisir à Paul.
"He cooked in order to please Paul."
 - (8) Il a déménagé aux Etats-Unis **après** avoir vécu 10 ans au Brésil.
"He moved to the United States after having lived in Brazil for 10 years."

On the contrary, in Dimlex, the prepositions that can take only noun as a argument are considered as connective, if this noun denotes an event, like *dank* in (9) :

- (9) **Dank** dieser Entwicklung kann die europäische Industrie den Anforderungen eines dynamischen Wettbewerbs gerecht werden

German prepositions that takes only nouns as arguments are listed in appendix. The French corresponding item do not occur in LexConn.

2.2 Pronominal adverbs

In German, many pronominal adverbs can be built on four different anaphoric forms (*da*, *wo*, *hier*), *so* + a preposition :

- *da-* + prep
- *wo-* + prep
- *hier-* + prep
- *so-* + prep

The direct translation in French would be a preposition with an anaphoric expression, be it a noun or a pronoun. See example (10) and its translation (11) :

(10) Am Montag stand ich drei Stunden in der Kaelte. **Dabei** habe ich mir einen kraeftigen Schnupfen zugezogen. (relation : circumstance)

(11) Lundi, je suis resté trois heures dans le froid. **A ce moment-là**, j'ai attrapé un gros rhume.

German so-called pronominal adverbs are listed in appendix. The French direct translations would be analysed more as AltLex (Alternative Lexicalisation, PTDB) than discourse connectives.

2.3 Preposition + infinitival clauses

In LexConn 30 prepositions are listed as connectives because they can take an infinitival clause as a complement. In German, only *um* (*in order to*), *ohne* (*without*) and *(an)statt* (*instead of*) can take an infinitival clause as an argument.

For example, on the one hand, you have a correspondance between the pair of preposition/subordinating conjunction *sans/sans que* in French (12) and the pair *ohne/ohne dass* in German (13) :

(12) a. Elle a trouvé un travail **sans** avoir d'aide.
b. Elle a trouvé un travail **sans qu'**on l'aide.
'She found a job, **without** getting any help.'

(13) a. Sie hat eine Arbeitstelle gefunden, **ohne** Hilfe zu haben.
b. Sie hat eine Arbeitstelle gefunden, **ohne dass** man ihr hilft.

On the other hand, the pair *avant de/avant que* in French (14) has no corresponding pair in German. The only possible way is to use the subordinating conjunction, as shown in (15) :

(14) a. Marie prend une douche **avant d'**aller au lit.
'Mary takes a shower before going to bed.'
b. Marie prend une douche **avant qu'**elle se couche.
'Mary takes a shower before she goes to bed.'
(unpreferred version)

(15) a. Marie nimmt eine Dusche **bevor** sie ins Bett geht.
b. *Marie nimmt eine Dusche **bevor** ins Bett zu gehen.

The list of pairs of preposition/subordinating conjunction in LexConn is listed in appendix.

2.4 PP or NP

Finally, LexConn contains many lexicalized NPs or PPs used as adverbial connectives. There are two kinds of such NPs or PPs in LexConn: (i) The ones that are lexicalised enough to be translated by another lexicalised PP in German (16) or by a item of another category (17) :

(16) a. Il est très serviable. **Par exemple**, il m'a aidé à déménager.
b. Er ist sehr hilfsbereit. **Zum Beispiel** hat er mir geholfen, um umzuziehen.
'He's very helpful. For example, he helped me to move.'

(17) a. Il est allé à Bruxelles, **au lieu d'**aller à Paris.
b. Er ist nach Bruxelles gegangen, **statt** nach Paris zu gehen.
'He went to Bruxelles, **instead of** going to Paris.'

But LexConn contains many more or less lexicalised PPs (e.g. *à l'instant où/at the instant when*) that wouldn't have any corresponding item in DimLex. Indeed, the German direct translation would be pure compositional PPs :

(18) a. La foudre est tombée sur la maison **à l'instant où** Marie est entrée.
b. Der Blitz schlug das Haus zum Augenblick, in dem Maria eintrat.
'The lightning struck the house at the moment when Mary came in.'

Form a general point of view, we expect that the difference between compositional and non-compositional expressions is idiosyncratic to each language, except for some expressions like *for example*.

3 Conclusion

More finegrained mismatches can be found between the two lexicons. It should be done together with the finalization of the (French-German) bilingual list in the future. Promissing collarborative works are planned for the next year.

A List of prepositions in DimLex

k9	angesichts Angesichts
k10	anhand Anhand an Hand An Hand
k11	anlässlich anläßlich anlaesslich anlaeßlich Anlässlich Anläßlich Anlaesslich Anlaeßlich
k19	aufgrund auf Grund Aufgrund Auf Grund
k42	dank (grâce à) Dank
k69	entgegen Entgegen
k76	für (pour) fuer Für Fuer
k77	gemäß gemaef Gemäß Gemaef
k91	infolge Infolge
k100	kraft Kraft
k101	laut Laut
k102	

mangels
Mangels
k121
qua
Qua
k141
statt
Statt
k142
trotz
Trotz
k151
vermöge
vermoege
Vermöge
Vermoege
k160
wegen
Wegen
k178
zufolge
Zufolge

B List of pronominal adverbials

k31

dabei
Dabei

k32

dadurch
Dadurch

k35

dafür
dafuer
Dafür
Dafuer

k38

dagegen
Dagegen

k44

darauf
Darauf
drauf
Drauf

k45

daraufhin
Daraufhin

k46

darob
Darob

k47

darum
Darum

k48

darüber hinaus
Darüber hinaus
darueber hinaus
Darueber hinaus

k51

dazu
Dazu

k53

demgegenüber
Demgegenüber
demgegenueber
Demgegenueber

k54

demgemäß
demgemaef
Demgemäß
Demgemaef

k65

drum

Drum
k223 davor
Davor
k81 hierauf
Hierauf
k82 hieraufhin
Hieraufhin
k83 hierdurch
Hierdurch
k84 hierfür
hierfuer
Hierfür
Hierfuer
k85 hiermit
Hiermit
k86 hiernach
Hiernach
k87 hierzu
Hierzu
k88 hingegen
Hingegen
k130 sodann
Sodann
k136 sonach
Sonach
k158 warum
k170 wonach
Wonach
k171 worauf
Worauf
k223 davor
Davor
k285 wo
Wo

k287

wodurch
Wodurch

C French prepositions followed by infinitival clauses

Table 1: Pairs of Prep/SubConj in LexConn and the corresponding entries in DiMLex (if any) or their possible translations in German (put in parenthesis)

Prep	SubConj
afin de	afin que
um, ... zu	damit
au lieu de	au lieu que
(an)statt,... zu	(an)statt, dass
plutôt que de	plutôt que
(an)statt,... zu	(an)statt, dass
avant de	avant que
	bevor
après	après que
	nachdem
jusqu'à	jusqu'à ce que
	bis
de façon à	de façon à ce que
	sodass
de manière à	de manière à ce que
	sodass
quitte à	quitte à ce que
(auf die Gefahr hin,... zu)	(auf die Gefahr hin, dass)
le temps de	le temps que
	(bis/während)
à moins de	à moins que
	(wenn/soweit/sofern... nicht)
à condition de	à condition que
(unter der Voraussetzung,... zu)	(unter der Voraussetzung, dass)
à seule fin de	à seule fin que
(mit dem einzigen Zweck, ... zu)	(mit dem einzigen Zweck, dass)
au moment de	au moment où
	(zum Zeitpunkt, wo / als)
dans l'espoir de	dans l'espoir que
(in der Hoffnung,... zu)	(in der Hoffnung, dass)
dans l'intention de	dans l'intention que
(in der Absicht,... zu)	(in der Absicht, dass)
dans le but de	dans le but que
(mit dem Ziel,... zu)	(mit dem Ziel, dass)