A TYPOLOGICAL *RARUM* IN SOGDIAN: OVERT COMPLEMENTIZERS IN INDICATIVE ROOT CLAUSES[1]

LUDOVICO FRANCO

ABSTRACT: Ludovico Franco. *A Typological* rarum *in Sogdian: Overt Complementizers in Indicative Root Clauses*. Lingua Posnaniensis, vol. LV (1)/2013. The Poznań Society for the Advancement of the Arts and Sciences. PL ISSN 0079-4740, ISBN 978-83-7654-273-7, pp. 55–67.

The purpose of this work is to illustrate an extremely rare linguistic feature, namely the overt present of a *root* complementizer in assertive/indicative (i.e. unmarked) matrix clauses, of the Sogdian language, an Eastern Middle Iranian Language once spoken in a region located in the valley of rivers Zaravshan and Kashkadarya (roughly corresponding to the territory of modern day Uzbekistan and Tajikistan).
This linguistic fact is very interesting because it represents an overt evidence of the principle of endocentricity *inferred* in the Generative tradition since the early 80s.
In comparative perspective, this uncommon feature of the Sogdian language may be associated to the mechanism of para-hypotaxis, previously studied in many different Romance languages (e.g. Old French, Old Italian, Old Catalan) and recently discovered in other genetically unrelated languages (e.g. Swahili, Zamucoan languages).

Ludovico Franco, Università degli Studi di Firenze, Italia, franco.ludovico@gmail.com

It is a trivial observation that root clauses with neutral speech act values, i.e. in the indicative/assertive mood, normally display no overt complementizers. Nevertheless, when matrix clauses "*serve to express questions, commands, exclamations or wishes, complementizers surface in a large number of languages*" (KENESEI & ORTIZ DE URBINA 1994: 7). Thus, it seems reasonable to suppose that the *markedness* of (non-neutral) root clauses can trigger the overt phonological realization of the head of the CP. Below in (1) you can see four examples of overt realization of the complementizer with interrogative forms (yes/no questions).[2]

(1a)   *Che* la   mangia   la   bimba   la   zuppa?   *Tuscan Italian*
           COMP SUBJ.CL.3SG.F eat.3SG.PRS the girl   the   soup
           'Does the child eat soup?'

---

[1] I wish to thank Luigi Rizzi and Rita Manzini for comments on the ideas developed in this paper. Many thanks to an anonymous reviewer for useful comments and suggestions which helped to improve this work. All errors are mine.

[2] HEINE & KUTEVA (2005: 56–58) consider this linguistic fact as the expression of a pathway of grammaticalization across conceptual domains, namely from *complementizer* to *polar question marker*. Indeed, cross-linguistically, there seems to be a path by which the use of a complementizers (especially those introducing indirect questions e.g. 'if,' 'whether') is extended to mark also yes/no questions.

(1b)   *Que*   vols   més   patates? *Catalan*
         COMP   want.2SG.PRS more   potatoes
         'Do you want more potatoes?' (HUALDE 1992: 2)

(1c)   *Vai*   viņ-š   jau   (ir)   at-nāc-is?   *Latvian*
         COMP   he-NOM.SG already COP.3SG.PRES here-come.PPA.NOM.SG.M
         'Has he come here already?' (STOLZ 1991: 67)

(1d)   *Aya*   in   gorbe-ye   šoma-st? *Persian*
         comp   this   cat-EZ   you-is
         'Is this your cat?' (MAHOOTIAN 1997: 9)

In the classic generative tradition, authors have *inferred* the presence of a (covert/null) complementizer head in (every) root clause by describing it as the landing site of head-movement (cf. TRAVIS 1984; CHOMSKY 1995; ROBERTS 2000; MATUSHANSKY 2006 among many others). A well-known example is given by structures containing *wh*-pronouns and other fronted constituents (e.g. modals) as shown in (2).

(2a)   What can you drink?

(2b)   [CP what$_i$ [C can$_j$ [TP you t$_j$ [VP drink t$_i$]]]]

Current researches in the minimalist framework, *enhance* the mechanism of feature strength as forcing or prohibiting the movement of an element to the C head (cf. ADGER 2003 and RADFORD 2004 among others) or argue from *that-trace* effects for the need of linking T to C head-movement with abstract Case in syntax (cf. PESETSKY & TORREGO 2001; BOECKX 2003; GALLEGO 2010). In the work that started the minimalist framework, CHOMSKY's (1995: 294) original proposal was that the C-head of matrix clauses is an *affix*. However, the difference between embedded (usually morphologically realized) and root CP (as said above usually null/covert) is quite difficult to capture along this line of reasoning.

It is relevant to say here that many recent works have questioned the existence of head movement in syntax, admitting only one type of movement (the phrasal one) and reanalyzing the head-type as an instance of remnant movement (cf. among others KOOPMAN & SZABOLCSI 2000; NILSEN 2003; KAYNE 2004; CINQUE 2005, 2010).[3] Also, it is worth noticing here that RIZZI (1997) has proposed very influential arguments for the decomposition of the complementizer phrase in a layered field with (at least) two poles, respectively Force, the higher one which is the interface to the higher context (e.g. the discourse, enhancing a window between syntax and pragmatic) and Finiteness, the lower one, which hosts non-finite Complementizers (like for examples Italian *di*, or English *for*) and acts as an interface to the propositional IP/TP.

Independently from the fine-grained (and competing) hypotheses roughly sketched above, crucial for the present work is the fact that a key point of the generative tradition

---

[3]   Remnant movement can be defined as the movement of an XP α from which extraction of β has taken place earlier in the derivation (cf. e.g. MÜLLER 1998) as shown in (i):

   (i)   [$_{\beta2}$ ... t1 ...] ... [... $_{\alpha1}$ ... [... t2 ...]]

(starting at least from the Government and Binding paradigm of research in the 1980s, cf. CHOMSKY 1981 and REINHART 1976, 1981) is the principle of *endocentricity* (namely, the generalization that phrases must have an *overt* or *covert* head).[4]

Hence, the existence of a language overtly realizing the complementizer head in an *unmarked* matrix clause could be seen as an *empirical* validation of an almost universal *postulation* (inferred e.g. by movement in generative syntax) of contemporary linguistic analysis.

Such a language, to my knowledge, has still not been clearly reported in the literature[5] and the scope of this paper is precisely to describe some noteworthy syntactic aspect of Sogdian, a Middle Eastern Iranian language, in which a complementizer was realized in root clauses with "greater than chance frequency".[6]

The paper is organized as follows. First, I will introduce the features of Sogdian relevant for the present discussion. Second, I will show that Sogdian used the same item for marking complementation / subordination and coordination and I will discuss this fact in cross-linguistic perspective. Third, I will describe the linguistic mechanism of para-hypotaxis, showing that, a related *device* can be, in principle, at work in Sogdian. The conclusion follows.

[4]   This principle is not exclusive of *orthodox* Generative Grammar and is accepted by other competing theories such as, for instance Word Grammar (HUDSON 1984), with possibly the sole exception of the Role and Reference Grammar paradigm (see VAN VALIN & LAPOLLA 1997).

[5]   Not uncontroversial claims for the existence of "declarative complementizers" have been already made in the literature. In previous work, LECARME (1999) focussed on the *pivotal role* of the 'assertive' C node in Somali at the interface levels, as point of contact between clause and discourse. Lecarme examined the distribution of the elements that are stardardly described as *focus markers* (the particles *baa/waa*) in Somali, showing that their distribution cannot be (fully) accounted for in *barely* functional terms. This fact is taken as evidence that these particles are not discourse markers at all and thus she argued that the "focus markers" of Somali are overt 'root' complementizers. MUGHAZY (2003: 1149) has shown that in Egyptian Arabic there is a pragmatic constraint (i.e. a marked context) that determines the presence of over complementizers in matrix clauses. In Egyptian Arabic overt complementizers such as *huwwa* 'he' and *da* 'that' are usually unacceptable except for 'focus structures' in which a speaker emphasizes the novelty of the information provided in the sentence and in *metalinguistic negation*, which is a specialized use of the negative operator where it functions as a device for registering an objection to a preceding utterance on any grounds other than its truth-conditional content (cf. HORN 1985). ETXEPARE (2010) shows that matrix clauses in Iberian Spanish optionally include a root complementizer (*que*). The presence of the complementizer adds a *reportative* component to the meaning of these clauses and the author analyzes these contexts as instances of (marked) quotative constructions. Finally the language in which the more unambiguous root-complementizer (*qu'; que*) seems to be attested is Gascon, spoken in southwestern France, where a second position (Wackernagel-like) particle usually follows the subject in the indicative mood and thus seems to instantiate indicative complementizer in root clauses (cf. HETZRON 1977; CAMPOS 1992). Consider the example below in (i):

(i)   lo Napoleon *qu'*   a   hèit hòrt un bon   ahar ad aqueth temps
        the N.   COMP has made strong a good affair at that   time
        'at that time Napoleon has made a pretty good deal' (PUSCH 2000: 189)

However this marker has been analysed as a (modal) particle signalling assertivity/evidentiality by some scholars (see e.g. FIELD 1985; PUSCH 2000, 2003).

[6]   Just to give an example, 90% ca. of the root clauses in the Sogdian *Ancient Letter V* are introduced by a 'matrix complementizer' (cf. GRENET et al. 1998: 92–93).

UBIQUITOUS COMPLEMENTIZERS IN SOGDIAN: AN OVERVIEW

The Sogdian language was an Eastern Middle Iranian SOV language spoken in the region of Sogdiana[7], currently located between Southern Uzbekistan and Western Tajikistan (see GERSHEVITCH 1954 and HESTON 1976 for detailed grammatical descriptions of Sogdian).[8]

Here, I will investigate a notable feature of Sogdian grammar, namely the presence in every clause (roots and subordinates) of an enclitic (or rarely free standing) complementizer (*əti*, *-ti*), encoded in a Wackernagel-like second position, to which other enclitics can be added (cf. SIMS-WILLIAMS 1989; QĀDERĪ 1996: 312; YAKUBOVICH 2002, 2005; YOSHIDA 2009: 314–315).[9] YOSHIDA (2009) was the first to identify *əti*, *-ti* as a 'universal' complementizer but he did not further explore the theoretical and typological implications of such very uncommon feature.[10] Consider the examples below in (3a, b). The same particle appears in second position both in the indicative (3a) and the interrogative (3b) sentence. The standard second position of *əti/-ti* in Sogdian seems to reveal that a root complementizer needs obligatory filled specifiers to be phonologically realized (cf. the similar behaviour of Gascon described in fn. 5). Note that in (3a) the same particle signals the adverbial clause expressing cause acting as a subordinator.

(3a)   čan šē satu əfsānx zāy əti   āγatim   paraw mā0-əti wānō patīγōš
       from 3 100 mile   land COMP came1SG.PST because-COMP thus hear.1SG.PST
       'I have come from the place 300 miles far (from here), because 1 heard thus…'
                                                                            (YOSHIDA 2009: 320)

(3b)   xa əspiyi-*ti*   kutsār wāčām
       the horses-COMP   where   I.send
       'Where shall I send the horses?'   (YOSHIDA 2009: 317)

Hence, we may assume a tendency toward a structure roughly as the one represented below in (4) for matrix clauses in Sogdian.

(4)   [*filled* SpecCP [C *əti/-ti* [TP [VP]]]]

---

[7]   Sogdians were Iranian in language and culture even if many features of their history and practice still remain mysterious (BENVENISTE 1929; MACKENZIE 1976; SIMS-WILLIAMS 1985; GRENET & SIMS-WILLIAMS 1987; SIMS-WILLIAMS 1989; YOSHIDA 2009). The presence of such a population has been recognized for a long time, by virtue of Chinese sources (cf. COMPARETI 2002). But it was only at the beginning of the 20th century that European archeologists recovered a relatively large number of documents written in the Sogdian language in Chinese Turkestan (the Xinjiang Uighur province). Then, soviet archeologists discovered many other documents in *proper* Sogdiana and new inscriptions were found along the *Silk Road* (COMPARETI 2002).

[8]   The Sogdian language was first documented around the 4th century AD and was attested until the 13th century (YOSHIDA 2006: 82, 2009: 329). Scholars argue that Sogdian dialects directly linked to ancient Sogdian are still spoken along the Yaghnobi River, in Tajikistan (cf. SIMS-WILLIAMS 1982: 69–70, YOSHIDA 2009: 327).

[9]   The presence of a subordinating particle after the first syntactic unit is fairly consistent. As shown in HESTON (1976: 281) with a set of relevant examples it appears: a) after a verb in the imperative; b) after a prepositional phrase; c) after an initial personal pronoun or after an initial personal pronoun preceded by a vocative marker; d) after negating, adverbial, interrogative or conditional particles; e) (less frequently) after nouns and adjectives.

[10]   Notice that possibly a similar behaviour can be assumed for Sogdian's cognate language Bactrian (cf. YAKUBOVICH 2005).

This fact is particularly interesting because many works in the Generative tradition since CHOMSKY & LASNIK (1977) have assumed a Doubly-Filled-Complementizer Filter of the type *[CP WH that].[11] Thus, in current terms complementizer deletion must obtain to satisfy this filter. In Sogdian C deletion clearly does not apply. Notice however that Sogdian *Wackernagel-like* complementizer is not without exceptions, especially for what concerns dependent clauses.[12]

Commonly, the overt specifier position of the Wackernagel-like complementizer of Sogdian matrix clauses is filled by the adverb (*ə*)*r-* (meaning 'also', see YOSHIDA 2009: 314), giving rise to the compound (*ə*)*r-ti* in the CP. Consider the examples below:

(5a)   *r-ti*   paδka xači əti   əxu zātē əwən   əptre   anδēk
       also-COMP law   is   COMP the   son the.GEN-DAT father.GEN-DAT manner
       δāre
       have.3SG.OPT
       'it is customary that a son should possess his father's manner'
                                                          (YOSHIDA 2009: 307)

(5b)   *ər-ti*   sāt   wispu ark   γərβām
       also-COMP whole all   work know1sg.prs
       'I know every kind of work'   (YOSHIDA 2009: 308)

Note again that in (5a) *əti* is also used as a marker of subordination (in a cleft structure), so that the matrix clause appears to be 'sandwiched' between two (identic) complementizers.

Interestingly the Sogdian root complementizers are overtly realized independently of the voice (and the position of the verb in the clause) and tense/aspect markers involved, as shown in (6).

(6a)   *r-ti*   xa zāy   zpart kira
       also-COMP the ground pure   was.made
       'the ground was made clean'. *Passive – verb final*   (HESTON 1976: 162)

(6b)   *r-ti*   kira   aβd   wārik   ēw zēn-γōδ
       also-COMP was.made seven lambskin one saddlecloath
       'one saddlecloath was made from seven lambskins'. *Passive – 'raised' verb*
                                                          (HESTON 1976: 162)

(6c)   *r-ti*   xu nawē   nγiγōδan pat[ī]menč
       also-COMP he new   garment wear.3SG.IMPFV
       'he wore a new garment'. *Imperfect*   (YOSHIDA 2009: 308)

---

[11] Nevertheless, very many countexamples are known (also) from Germanic (e.g. Bavarian, BAYER 1984 or West Flemish, HAEGEMAN 1992) or Romance languages (e.g. North Italian dialects, see BENINCÀ 2001 and references cited there, among many others). Consider also BALTIN (2010) for a recent proposal concerning this point (namely, against the existence of doubly filled complementizers due to a far more layered structure in the C field).

[12] For instance, in Sogdian, purpose clauses are linked to the matrix clause by the complementizer *əti* without any other particle/word preceding it (see HESTON 1976, ch. 4). See also YAKUBOVICH (2005: 211, 216) for a schematic representation of *əti* distribution. Further notice that a *quasi* Wackernagel position for clitics, which optionally can float within clauses, seems to be an ubiquitous feature in present day Eastern Iranian languages (ERSHLER 2010). For a formal treatment of the Wackernagel law, perfectly fitting Sogdian data see AGBAYANI et al. (2010).

(6d) pār-*ti*-šī   xu wīnā žani frāmātδāre
      but/for-COMP-him the lute to.play order.2SG.PST
      'but you ordered him to play the lute'. *Preterite*   (YOSHIDA 2009: 315)

(6e) rti   əču   wānō   zārī   rāy-e-əskun
      also-COMP why thus   miserably cry2SG.PRS-DUR
      'why are you crying so sorrowfully?' *Progressive*   (YOSHIDA 2009: 308)

Interestingly, also imperative and negative forms can display the overt root complementizer. Consider the examples in (7).

(7a) əti xēp0 βrāt   par žəγma nā   ōzām
      COMP own brother on lie   PROHIBITIVE condemn.2SG.IMPER
      'do not condemn your brother with lies'. *Imperative*   (YOSHIDA 2009: 311)

(7b) r-*ti*-βi   nē δūrī zāy šaw-t
      also-COMP-ENCL   not far land go.3SG.PRS
      'he did not go far'. *Negative*   (HESTON 1976: 204)

Specifically, the fact that complementizers are present together with imperative is quite interesting because many contemporary accounts of imperatives involve the C head or the CP field (see e.g. HAN 2000; ZANUTTINI 2008; cf. also MANZINI & SAVOIA 2008 for an alternative implementation). Sogdian data seem to go against such views (the C slot is filled) or, alternatively, support a *layered* C field à la RIZZI (1997) (*contra* a unitary CP).[13]

Another interesting fact is that a particle of the same form of the root one is obligatorily found in Sogdian relative clauses, in which relative pronouns (delimiting clausal edges) are invariantly followed by an overt complementizer: relative pronouns normally combine with the complementizer *əti*, *-ti*, leading to compound sequences such as *kē əti*, *kē-ti* "who, which" *ču əti*, *ču-ti* "which," *ku əti* "where," *čānō əti* "how" (SIMS-WILLIAMS 1989; SKJÆRVØ 2007). See the examples below in (8), retrieved from Classical Sogdian and taken from YOSHIDA 2009: 318, cf. also FRANCO 2012).[14]

(8a) …ōnō martī wiru   kunāt   ke-*ti*-šī   xwati   rēžāt
      that man   husband make.3SG.SUBJ REL-COMP-her herself please.3SG.SUBJ
      'she shall make that man her husband who might be pleasing to her'

(8b) …əwən šē   ratne… əkya-*at*-mī   sāče   xu pāš   pāt
      the   three   jewel… REL-COMP-me   be-fitting 3SG.OPT. the honour observe INF
      'three jewels … to which it may be fitting for me to pay honour'

(8c) …xānd āfrītēt ōtākt … ku-*əti*   waδēδ mēnand xa roxšnda βaγīšt
      those blessed places … where(REL)-COMP there stay3PL the light   gods
      'those blessed places where the light gods are staying'

The pervasiveness of the feature under consideration and the fact the root complemen-

---

[13] Nevertheless, regarding this point, Sogdian (which notably allow an imperative verb to precede the Wackernagel complementizer; cf. HESTON 1976) can give only *hints*, due to impossibility of eliciting fine-grained grammatical 'live' judgements (e.g. topic licensing, wh-extraction, optionality of constructions, etc.).

[14] Interestingly Sogdian relative clauses, as shown in (8), are invariantly signalled by a *cataphoric* demonstrative pronoun (e.g *ōnō*) in initial position in the matrix clause (PROVASI 1997).

tizer is identic with the subordinate complementizer lead us to reject the possibility that *əti, -ti* may be a (mere) marker of illocutionary force. Furthermore, from a diachronic point of view, the overt root complementizer is present in every stage of the Sogdian language, starting from the oldest extant Sogdian texts, the so-called *Ancient Letters* and being ubiquitously present in Buddhist, Manichaean and (late) Christian texts.[15] Consider the examples below (adapted from WENDTLAND 2011, in which complete references to the documents and to the sources are provided).

(9a)   ꞽḤR-*ZY*   xwtꞽynβ   ꞽst   8   srδ   ꞽYK-*ZY*=m   pryšt   cntry   sꞽr
         also-COMP   sirs   be.3SG   8   year   REL-COMP-me   send.PST-PTC   inside   to
         sꞽγrꞽk   ꞽPZY   prnᵐγt
         Saghrak   and   Farn-āghat
         'And, sirs, it is eight years since I sent Saghrak and Farn-āghat inside'[16]
                                       (*Ancient Letter II*, 31-32; SIMS-WILLIAMS 2001: 271)

(9b)   ꞽzw   *ZY*   tꞽβꞽkH   sꞽr   mzꞽyx   ꞽrn   *ZY*   γwꞽnH   ꞽkrtw   δꞽrꞽm
         I   COMP   you.ACC   towards   great   fault   COMP   sin   make.PST-PTC   have.1SG
         'I have committed great faults and sin(s) against you'
                                       (*Vessantara Jātaka 1452*, [Buddhist text] BENVENISTE 1946: 83)

(9c)   r-*ti*   ᵐγšδꞽrt   ꞽkw   ꞽBYw   sꞽr   ZKn   krn   kꞽt   zꞽky   γzꞽt
         also-COMP   began   to   father   to   the.GEN-DAT   deaf   dumb   boy   slander
         'and she began to slander the deaf-dumb boy to his father'
                                       (*KG 2, 20*, [Manichean text] SIMS-WILLIAMS 1990: 284)

(9d)   ꞽt   pw   dbn   žwꞽdꞽrt   cꞽfyd   pstnꞽ   byrdꞽrt
         COMP   without   fear   live.HAVE.3SG.PERF   in.as.much   respite   find.HAVE. 3SG.PERF
         'And has lived without fear, in as much as he has found respite!'
                                       (*manuscript C2, 51 r 26-27*, [Christian Text] SIMS-WILLIAMS 1985: 91)

COORDINATORS, SUBORDINATORS AND *ROOT* COMPLEMENTIZERS

Very interestingly, the Sogdian particle *ti/əti* (cf. YAKUBOVICH 2005: 201) functions also as a subordinating particle as shown in the examples in (8) above, and as a coordinating item (both at phrasal and clausal level), as shown in (9a). Consider the example reported below in (10) where we may observe again the presence of the same marker employed as a root complementizer and as a coordinating linker at the phrasal level.

(10)   r-*ti*   ō   barāman   xwart   *əti*   čašant   xwer
         also-COMP   the.ACC   Brahmin   food   and   drink   feed.IMPF.3s
         'he made the Brahmin eat (fed him with) food and drink' (YOSHIDA 2009: 303)

---

[15]   Notice however that the typological *rarum* represented by the over matrix complementizer tended to disappear in late (Christian) Sogdian manuscripts in which, when present, it was expressed often by the contract form 't (see YOSHIDA 2009; YAKUBOVICH 2005: 216).

[16]   Cf. YAKUBOVICH 2005 for a description of the *path* that had lead from the Aramaic to the Sogdian *heterogram* ZY (corresponding to the complementizer as illustrated in the previous examples; for an concise introduction to Sogdian script(s) see GHARIB 1995: xxvii–xxix).

It is also worth noticing that – like the subordinator in the relative clauses introduced in (8) – the coordinating linker is used in 'compound' forms (cf. SIMS-WILLIAMS 1989: 191) to express adversative coordination as in example (6b) above and disjunction as shown below in (11).

(11)   ču-*ti*   xa   zākt   taδēδ   aɣatant   katār-*ti*   ne
       whether-COMP the   children there   came   or-COMP   not
       'Have the children arrived to you or not?'   (YOSHIDA 2009: 317)

The identity of what YAKUBOVICH (2005) calls the 'demarcational' *ti*/ə*ti* in root clauses, the *standard* subordinating complementizer and the '*and*' conjunction is a very intriguing feature.

Subordinators are often contrasted with coordinators, but albeit this distinction works quite well with widely studied languages (e.g. Germanic, Slavic or Romance), it becomes more problematic from a typological point of view. For instance, in so-called *clause-chaining* languages (e.g. Papuan and Amerindian), adverbial subordinators seem to *overlap* with coordinators as shown in LONGACRE (2007: 364, 375ff; cf. also NORDSTRÖM 2010: 97, HASPELMATH 2004). But, the lack of clear distinction between subordinators and coordinating linkers is quite widespread elsewhere. Just to give an example, consider the data below from Fongbe, a Kwa language spoken in Benin.

(10a)   Kɔ̀kú   wá   bɔ̀   Àsíbá   yì
        Koku   arrive   CONJ   Asiba   leave
        'Koku arrived and-then Asiba left'   (LEFEBVRE 2004: 125)

(10b)   Kɔ̀kú   kò   yì   có   bɔ̀   à   (ná)   wá
        Koku   ANT   leave   before CONJ   2sg   DEF.FUT arrive
        'Koku had left before you arrived'   (LEFEBVRE 2004: 132)

Furthermore, CULICOVER & JACKENDOFF (1997) has shown that there are many instances of coordination-subordination mismatches even in very well studied languages (cf. also YUASA et al. 2002). So according to them, *'left constituent-subordinating' and-constructions* in English as in (11) should be asymmetrically characterized in the realm of syntax *vs.* semantics, being coordinate in the former, and subordinate in the latter. In other words, the conjunction in (11) should be interpreted like the conditional subordinator in (12).

(11)   You drink one more can of beer *and* I'm leaving.

(12)   *If* you drink one more can of beer I leave.

The facts roughly sketched in the present section allow us to introduce the proposal of a fairly consistent para-hypotactic strategy employed by Sogdian at clausal level.

PARA-HYPOTAXIS IN SOGDIAN?

The name *para-hypotaxis* has been first employed by SORRENTO (1929) to designate sentences of the type [SUB – *dependent clause* – COORD – *matrix clause*], namely with a proleptic dependent clause, and the main clause preceded by a coordinator. The term is still widely employed by Romance linguists (see BERTINETTO & CIUCCI 2012; MAZZOLENI 2002; DE CAPRIO 2010). Para-hypotaxis was widespead in Romance (e.g. Old French, Old Italian,

Old Catalan, Old Occitan, etc.) until the XV century. Examples of Romance para-hypotactic structures (taken from BERTINETTO & CIUCCI 2012: 90) are reported below in (13).

(13a)　　*E　si*　venoi-lh　Frances,　*que*　vulhan　asautar,　*e*　nos　ab
　　　　　　and　if　come.3PL　French　who　want.3PL　assault　and　we　with
　　　　　　las　balestas　es　far-em　tot-z　nafrar
　　　　　　the　arbalests　them　do-FUT.3PL　all-PL　wound.INF
　　　　　　'And if the French people come and want to assault us, we shall wound them with the arbalests'
　　　　　　Old Occitan (Chanson de la croisade albigeoise, Bartsch-Koschwitz, Chres-tomathie Provençale, 1904: 203)

(13b)　　*S'*　io　dissi　il　falso,　*e*　tu　falsasti　il　conio
　　　　　　If　I　say.1SG.PST　the　false　and　you　alter-2SG.PST　the　minting.die
　　　　　　'If I said something false, you (did worse, for you) altered the minting die'
　　　　　　Old Italian (Dante Alighieri, Commedia, Inf., 30.115)

This kind of structures has been interpreted as an areal (i.e. Old Romance) and archaic feature. Contemporary research has however shown extensive use of para-hypotactic constructions in modern languages (BERTINETTO & CIUCCI 2012: 91). For instance, REBUSCHI (2001) has shown that these constructions are present in Swahili and BERTINETTO & CIUCCI (2012) have convincingly shown that para-hypotaxis is a pervasive feature in the present day Zamucoan languages, spoken in Paraguay and Bolivia. Consider the examples below:

(14a)　　Mtu　ye　*yote*　akitaka　kunifuata　*na*　ajikane　mwenyewe
　　　　　　man　all　if　3SG.want　1SG.follow　and　3SG.deny.SUBJ　3SG.RFL
　　　　　　'Should anyone want to follow me, he should renounce to himself'
　　　　　　Swahili (REBUSCHI 2001, ex. [38], taken from BERTINETTO & CIUCCI 2012: fn. 3)

(14b)　　*Uje*　ye　t-uu_leeych,　*ich*　ese　aahn-t　s-erz　yoo
　　　　　　SUB　NEG　1SG-fight　COORD　DEM.MS　evil_spirit-MS　3-win　1sg
　　　　　　'When/if I don't fight, that evil spirit will defeat me'
　　　　　　Chamacoco, Zamucoan [CIUCCI, field-work], taken from BERTINETTO & CIUC-CI 2012: 98)

A para-hypotactic strategy must definitely be assumed for Sogdian. A subordinate clause that states a logical/temporal dependency usually precede the root clause, and is typically marked by conjunctions such as *kaδ* 'if', *čānō* 'when, since', *mant* 'when, while, until', *kū parm* 'as long as, until', etc. (HESTON 1976: 216; YOSHIDA 2009: 319). Crucially, both subordinate and matrix clauses are 'signalled' (i.e. in the first stages of the language, documented in the *Ancient Letters*, cf. GHARIB 1995: xix) by the 'compound conjunction' *r-ti* (glossed elsewhere with 'also-COMP', glossed below with 'and' for brevity). To our knowledge para-hypotaxis has not been previously associated with Sogdian. Consider the examples below:

(15a)　　[[*rti*-šu　*ka*δ　nē　waγtu-δāre]　*rti*-šu　yōnēδ　na　wāč]
　　　　　　and-him　if　not　send-PRET.2GS　and-him　immediately　not　send.2SG.IMP
　　　　　　'if you have not sent him, do not send him immediately'

　　　　　　　　　　　　　　　　　　　　　　(adapted from YOSHIDA 2009: 319)

(15b)   [[*rti   čānō* əxu suδāšn waγtu-wan-a] *rti* əxu barāman žγart   manxaz]
         and   when the S.   said-POT   and   the brahmin quickly rose-I PF
         'when/after Sudhashn had said (it) the brahmin got up quickly'
                                                        (adapted from YOSHIDA 2009: 302)

(15c)   [[*rti-šu kū parm* nē   tōžān]   *rti*-šu… paru   wartu   δārām-kām]
         and-it   as.long.as not   pay.1SG.SUBJ and-it… on   interest have.1SG.FUT
         'until I pay it (= as long as I do not pay it), I shall owe (lit. "have") it with inter-
         est'   (adapted from YOSHIDA 2009: 319)

The notable fact of Sogdian is that the particle *r-ti* traditionally assumed as a marker of both coordination and subordination, appears in both the main and the subordinate clause. This case is particularly interesting from a theoretical viewpoint because it represent an evidence to recent generative approaches to para-hypotaxis, which assume that the 'co-ordinator' in the matrix clause of such constructions is actually a particle hosted in the (layered field of) the complementizer (see BENINCÀ 2001; POLETTO 2006).[17] If such an ap-proach is on the right track, we have further evidence for a *root* complementizer status of 'demarcational' *əti*, -*ti*. Sogdian – like many other languages – *underspecifies* (and/or as-sembles compositionally,[18] see the previous paragraphs) subordination and coordination in many constructions.[19] Hence, in Sogdian we have situations (previously undescribed in the literature, to my knowledge) in which a proleptic dependent clause and the root clause are introduced by the same item.

CONCLUSION

In this (mainly descriptive) paper we have illustrated a very rare feature of the Sogdian language, namely the presence of a ubiquitous complementizer particle in the root clause. This fact is quite important because it provides empirical evidence for the principle of *endo-centricity*, a hallmark in the generative approach to language. We have shown that Sogdian underspecifies between coordination and subordination and we have shown that this lin-guistic fact is not uncommon from a typological perspective. Finally, we have seen that the mechanisms of para-hypotaxis may be at work is the Sogdian language. This mechanism, from a theoretical viewpoint (cf. POLETTO 2006), may be seen as an evidence for the com-plementizer status of the particle *əti*, -*ti* in the matrix clauses. If we are on the right track, we have found that at least one Iranian language in which para-hypotaxis was at work. In future research we will try to see if this interesting syntactic phenomenon is present and/or has been explored in other Indo-Iranian languages.

[17] Specifically, POLETTO (2006: 232) argues that Italian *e* (and) conjunction in paratactic constructions is actually a *topic marker* in the layered CP and proposes a structure such the one represented below:

(i)   [$_{TopicP}$ (Null) Top [$_{Topic°}$ e [$_{TopicP}$ [$_{CP}$]]]]

[18] Compound 'complementizers' are quite common cross-linguistically. Just to give two examples consider the Persian *agar ke* (if), and the Italian *perché* (because), where, respectively, *ke* and *che* are the 'bare' comple-mentizers, used to introduce e.g. complement clauses.

[19] Consider YAKUBOVICH 2005 for a complete taxonomy that shows also the contexts in which Sogdian seems to discriminate among subordination and coordination; cf. also fn. 9.

REFERENCES

Adger David. 2003. *Core Syntax*. Oxford University Press: Oxford.

Agbayani Brian, Golston Chris. 2010. "Second-position is First-position: Wackernagel's Law and the Role of Clausal Conjunction." *Indogermanische Forschungen* 115, 1–21.

Ambrosini Riccardo, Bologna Maria Patrizia, Motta Filippo, Orlandi Chatia (eds.). 1997. *Scribthair a ainm n-ogaim. Scritti in memoria di Enrico Campanile*. Pisa: Pacini.

Andersen Gisle, Fretheim Thorstein (eds.). 2000. *Pragmatic Markers and Propositional Attitude*. Amsterdam: John Benjamins.

Baltin Mark. 2010. "The Non-reality of Doubly Filled Comps." *Linguistic Inquiry* 41, 331–335.

Baltin Mark, Collins Chris (eds.). 2000. *Handbook of Syntactic Theory*. Oxford: Blackwell.

Batllori Montserrat, Hernanz Maria Luisa, Picallo Carme, Roca Francesc (eds.). 2006. *Grammaticalization and Parametric Variation*. Oxford: Oxford University Press.

Bayer Josef. 1984. "COMP in Bavarian Syntax." *The Linguistic Review* 3, 209–274.

Belletti Adriana (ed.). 2004. *Structures and Beyond*. New York: Oxford University Press.

Benincà Paola. 2001. "The Position of Topic and Focus in the Left Periphery." In: Cinque & Salvi 2001: 39–64.

Benveniste Émile. 1929. *Essai de grammaire sogdienne. Deuxième partie: morphologie, syntaxe et glossaire*. Paris: Paul Geuthner.

Benveniste Émile. 1946. *Vessantara Jātaka [Mission Pelliot en Asie Centrale IV]*. Paris: Paul Geuthner.

Bertinetto Pier Marco, Ciucci Luca. 2012. "Parataxis, Hypotaxis and Para-Hypotaxis in the Zamucoan Languages." *Linguistic Discovery* 10, 89–111.

Boeckx Cedric. 2003. *Islands and Chains. Resumption as Stranding*. Amsterdam: John Benjamins.

Campos Héctor Ramiro. 1992. "Enunciative Elements in Romance." *Linguistics* 30, 911–940.

Chomsky Noam. 1981. *Lectures on Government and Binding*. Dordrecht: Foris.

Chomsky Noam. 1995. *The Minimalist Program*. Cambridge, Mass.: MIT Press.

Chomsky Noam, Lasnik Howard. 1977. "Filters and Control." *Linguistic Inquiry* 8, 425–504.

Cinque Guglielmo. 2005. "Deriving Greenberg's Universal 20 and Its Exceptions." *Linguistic Inquiry* 36, 315–332.

Cinque Guglielmo. 2010. *The Syntax of Adjectives. A Comparative Study*. Cambridge, Mass.: MIT Press.

Cinque Gugliemo, Salvi Giampaolo. 2001. *Current Studies in Italian Syntax. Essays Offered to Lorenzo Renzi*. Amsterdam: Elsevier North Holland.

Compareti Matteo. 2002. "Introduction to the History of Sogdiana." In: Güzel et al. 2002: 373–381.

Culicover Peter W., Jackendoff Ray. 1997. "Semantic Subordination Despite Syntactic Coordination." *Linguistic Inquiry* 28, 195–217.

De Caprio Chiara. 2010. "Paraipotassi e sì di ripresa. Bilancio degli studi e percorsi di ricerca (1929–2010)." *Lingua e Stile* XLV, 285–328.

Erschler David. 2010. "On Optionality in Grammar: The Case of East Iranian almost Wackernagel Clitics." *Hand-out of a Talk Given at Syntax of the World's Languages IV, Lyon, 2010*.

Etxepare Ricardo. 2010. "From Hearsay Evidentiality to Samesaying Relations." *Lingua* 120, 604–627.

Field Thomas T. 1985. "Speech Act Markers in Modern Gascon." In: King & Maley 1985: 77–97.

Franco Ludovico. 2012. "Against the Identity of Complementizers and (Demonstrative) Pronouns." *Poznań Studies in Contemporary Linguistics* 48, 565–596.

Gallego Ángel. 2010. *Phase Theory*. Amsterdam: John Benjamins.

Gershevitch Ilya. 1954. *A Grammar of Manichean Sogdian*. Oxford: Oxford University Press.

Gharib Badr-uz-zaman. 1995. *Sogdian Dictionary*. Tehran: Farhangan.

Gignoux Philippe. 1987. *Transition Periods in Iranian History: Actes du Symposium de Fribourg-en-Brisgau (22–24 Mai 1985)*. Leuven: Peeters.

Grenet Franz, Sims-Williams Nicholas. 1987. "The Historical Context of the Sogdian Ancient Letters." In: Gignoux 1987: 101–122.

Grenet Franz, Sims-Williams Nicholas, de la Vaissière Étienne. 1998. "The Sogdian Ancient Letter V." *Bulletin of the Asia Institute* 12, 91–104.

Güzel Hasan Celâl, Oğuz Cem, Karatay Osman (eds.). 2002. *The Turks*. Vol. 1. *Early Ages*. Ankara: Yeni Türkiye.

Haegeman Liliane. 1992. *Theory and Description in Generative Syntax: A Case Study in West Flemish.* Cambridge: Cambridge University Press.

Haegeman Liliane (ed.). 1997. *Elements of Grammar: Handbook of Generative Syntax.* Dordrecht: Kluwer.

Han Chung-Hye. 2000. *The Structure and Interpretation of Imperatives: Mood and Force in Universal Grammar.* New York: Garland.

Haspelmath Martin (ed.). 2004. *Coordinating Constructions.* Amsterdam: John Benjamins.

Heine Bernd, Kuteva Tania. 2005. *Language Contact and Grammatical Change.* Cambridge: Cambridge University Press.

Heston Wilma. 1976. "Selected Problems in Fifth to Tenth Century Iranian Syntax." Doctoral dissertation, University of Pennsylvania.

Hetzron Robert. 1977. "La particule énonciative qué en gascon." *Studi Italiani di Linguistica Teorica ed Applicata* 6, 161–221.

Horn Laurence R. 1985. "Metalinguistic Negation and Pragmatic Ambiguity." *Language* 61, 121–174.

Hualde José Ignacio. 1992. *Catalan.* London: Routledge.

Hudson Richard. 1984. *Word Grammar.* Oxford: Blackwell.

Kayne Richard S. 2004. "Prepositions as Probes." In: Belletti 2004: 192–212.

Kenesei Istvan, Ortiz de Urbina Jon. 1994. "Functional Categories in Complementation." [Accepted for publication in Complementation in the Languages of Europe, ed. by N. Vincent. Dordrecht: Kluwer]. Unpublished due to editor's failure.

Kenstowicz Michael (ed.). 2001. *Ken Hale. A Life in Language.* Cambridge, Mass: MIT Press.

King Larry D., Maley Catherine A. (eds.). 1985. *Selected Papers from the XIIIth Linguistic Symposium on Romance languages, Chapel Hill, N.C., 24–26 March 1983.* Amsterdam: John Benjamins.

Koopman Hilda, Szabolcsi Anna. 2000. *Verbal Complexes.* Cambridge, Mass: MIT Press.

Korn Agnes, Haig Geoffrey, Karimi Simin, Samvelian Pollet (eds.). 2011. *Topics in Iranian Linguistics.* Wiesbaden: Reichert Verlag.

Lecarme Jacqueline. 1999. "Focus in Somali." In: Rebuschi & Tuller 1999: 275–309.

Lefebvre Claire. 2004. "Coordinating Constructions in Fongbe with Reference to Haitian Creole." In: Haspelmath 2004: 123–164.

Longacre Robert E. 2007. "Sentences as Combination of Clauses." In: Shopen 2007: 372–420.

MacKenzie David N. 1976. *The Buddhist Sogdian Texts of the British Library.* Tehran–Liège: Brill.

Mahootian Shahrzad. 1997. *Persian.* London: Routledge.

Manzini M. Rita, Savoia Leonardo M. 2011. "Mesoclisis in the Imperative: Phonology, Morphology or Syntax?" *Lingua* 121: 1101–1120.

Matushansky Ora. 2006. "Head Movement in Linguistic Theory." *Linguistic Inquiry* 37, 69–109.

Mazzoleni Marco. 2002. "La 'paraipotassi' con ma nell'italiano antico: verso una tipologia sintattica della correlazione." *Verbum* IV, 399–427.

Mughazy Mustafa. 2003. "Metalinguistic Negation and Truth Functions: The Case of Egyptian Arabic." *Journal of Pragmatics* 35, 1143–1160.

Müller Gereon. 1998. *Incomplete Category Fronting.* Dordrecht: Kluwer.

Nilsen Øystein. 2003. *Eliminating Positions: The Syntax and Semantics of Sentence Modification.* Doctoral dissertation, University of Utrecht.

Nordström Jackie. 2010. *Modality and Subordination.* Amsterdam: John Benjamins.

Pesetsky David, Torrego Esther. 2001. "T-to-C Movement: Causes and Consequences." In: Kenstowicz 2001: 355–426.

Poletto Cecilia. 2006. "Sì and e as CP expletives in Old Italian." In: Batllori et al. 2006: 206–235.

Provasi Elio. 1997. "Note sulle costruzioni relative in Sogdiano." In: Ambrosini et al. 1997: 815–839.

Pusch Claus D. 2000. "The Attitudinal Meaning of Preverbal Markers in Gascon: Insights from the Analysis of Literary and Spoken Language Data." In: Andersen & Fretheim 2000: 189–206.

Pusch Claus D. 2003. "Preverbal Modal Particles in Gascony Occitan." *Belgian Journal of Linguistics* 16, 105–118.

Qāderī Teymūr. 1996. "Ğomalāt-e nā-hampāye dar zabān-e soγdī" [Hypotactic Sentences in Sogdian]. *Farhang* 17, 311–316.

Radford Andrew. 2004. *Minimalist Syntax: Exploring the Structure of English.* Cambridge: Cambridge University Press.

Rebuschi George, Tuller Laurice (eds.). 1999. *The Grammar of Focus.* Amsterdam: John Benjamins.

Rebuschi Georges. 2001. "Coordination et subordination. Première partie: la co-jonction restreinte." *Bulletin de la Société de Linguistique de Paris* 96, 23–60.

Reinhart Tanya. 1976. *The Syntactic Domain of Anaphora*. Doctoral dissertation, MIT.

Reinhart Tanya. 1981. "Definite NP anaphora and c-command domains." *Linguistic Inquiry* 12, 605–635.

Rizzi Luigi. 1997. "The Fine Structure of the Left Periphery." In: Haegeman 1997: 281–337.

Roberts Ian. 2000. "Head Movement." In: Baltin & Collins 2000: 113–147.

Schmidt Maria Gabriela, Bisang Walter (eds.). 2001. *Philologica et Linguistica. Historia, Pluralitas, Universitas. Festschrift für Helmut Humbach zum 80. Geburstag am 4. Dezember 2001.* Trier: Wissenschaftlicher Verlag.

Schmitt Rüdiger (ed.). 1989. *Compendium Linguarum Iranicarum.* Wiesbaden: Reichert.

Shirai Satoko, Shōgaito Masahiro. 2006. *Philological Studies on Old Central Asian Manuscripts.* Kyoto: Kyoto University.

Shopen Timothy 2007. *Language Typology and Syntactic Description.* Vol. II: *Complex Constructions* [2nd ed.]. Cambridge: Cambridge University Press.

Sims-Williams Nicholas. 1982. "The Double System of Nominal Inflexion in Sogdian." *Transactions of the Philological Society* 80, 67–76.

Sims-Williams Nicholas. 1985. *The Christian Sogdian Manuscript C2.* Berlin: Akademie-Verlag.

Sims-Williams Nicholas. 1989. "Sogdian." In: Schmitt 1989: 173–192.

Sims-Williams Nicholas. 1990. "The Sogdian Fragments of Leningrad II: Mani at the Court of Shahanshah." In: *Bulletin of the Asia Institute 4, In honor of Richard Nelson Frye. Aspects of Iranian Culture*, 281–288.

Sims-Williams Nicholas. 2001. "The Sogdian Ancient Letter II." In: Schmidt & Bisang 2001: 267–280.

Skjærvø Prods Oktor. 2007. "An Introduction To Manichean Sogdian." Ms. Harvard University.

Sorrento Luigi. 1929. "Il fenomeno della paraipotassi nelle lingue neolatine." *Rendiconti del Reale Istituto Lombardo di Scienze e Lettere* 52, 449–463.

Stolz Thomas. 1991. *Sprachbund im Baltikum? Estnisch und Lettisch im Zentrum einer sprachlichen Konvergenzlandschaft.* (Bochum-Essener Beiträge zur Sprachwandelforschung, 13). Bochum: N. Brockmeyer.

Travis Lisa. 1984. *Parameters and Effects of Word Order Variation.* Doctoral dissertation, MIT.

Van Valin Robert D., LaPolla Randy J. 1997. *Syntax.* Cambridge: Cambridge University Press.

Wendtland Antje. 2011. "The Emergence and Development of the Sogdian Perfect." In: Korn et al. 2011: 39–52.

Windfuhr Gernot (ed.). 2009. *The Iranian Languages.* London: Routledge.

Yakubovich Ilya. 2002. "Mugh 1.I. Revisited." *Studia Iranica* 31, 231–253.

Yakubovich Ilya. 2005. "The Syntactic Evolution of Aramaic ZY in Sogdian." *Studia Iranica* 34, 199–230.

Yoshida Yutaka. 2006. "Honorific and Polite Expressions in Sogdian." In: Shirai & Shogaito 2006: 81–94.

Yoshida Yutaka. 2009. "Sogdian." In: Windfuhr 2009: 279–334.

Yuasa Etsuyo, Sadock Jerrold M. 2002. "Pseudo-subordination: A Mismatch Between Syntax and Semantics." *Journal of Linguistics* 38, 87–111.

Zanuttini R. 2008. "Encoding the Addressee in the Syntax: Evidence from English Imperative Subjects." *Natural Language and Linguistic Theory* 26, 185–218.

Allatum die 6 mensis Januarii anno 2013