

## *Vertepe: una proposta di annotazione semantica XML-TEI della traduzione antica*

### *La creazione del primo corpus della traduzione antica<sup>1</sup>*

L'avvento della rivoluzione informatica ha profondamente segnato anche il panorama degli studi classici: negli ultimi decenni si è assistito alla nascita di banche dati della letteratura antica sia esclusivamente della letteratura latina<sup>2</sup> e della letteratura greca<sup>3</sup> sia della letteratura greco-latina<sup>4</sup>.

<sup>1</sup> Questo lavoro è stato realizzato all'interno del progetto *Le strategie del uertere in età tardoantica: una proposta di annotazione XML-TEI della traduzione*, finanziato dalla Fondazione CRT. Sono grata a Raffaella Tabacco per l'attenta rilettura e a Luigi Battezzato per i preziosi consigli. Desidero, infine, ringraziare i revisori anonimi per le indicazioni suggerite.

<sup>2</sup> Punti di riferimento per la libera consultazione della letteratura latina *online* sono le seguenti biblioteche digitali: il *Phi* per le opere letterarie dalle origini al II sec. d.C. (<https://latin.packhum.org>), *Musisque deoque* per la poesia (<http://www.mqdq.it>) e *digilibLT* per la prosa pagana di epoca tardoantica (<http://www.digiliblt.uniupo.it>), *alim* per i testi in latino dal XIII al XVII sec. (<http://alim.dfl.univr.it>); mentre per i testi cristiani la risorsa principale resta *LLT* consultabile a pagamento (<https://about.brepolis.net/library-of-latin-texts/>). Sulle biblioteche digitali si veda Lana 2012a.

<sup>3</sup> Fondamentale per la consultazione della letteratura greca è l'ampia banca dati *TLG* ora parzialmente ad accesso libero *online* (<http://stephanus.tlg.uci.edu>), a cui si affiancano biblioteche digitali dedicate ad uno specifico genere o autore, come il *corpus* degli scolii a Euripide (<https://euripidesscholia.org>).

<sup>4</sup> Emblematico è il caso di *Perseus* (<http://www.perseus.tufts.edu>), risorsa imprescindibile per la libera consultazione *online* della letteratura antica.

### *Una proposta di annotazione semantica*

Nel caso di traduzioni, i testi greci e latini vengono però inseriti in tali biblioteche digitali individualmente, in modo non sistematico, e senza possibilità di visualizzazione parallela, come avviene invece regolarmente nei *corpora* dedicati allo studio di traduzioni nell'ambito delle letterature moderne<sup>5</sup>, e occasionalmente anche in edizioni cartacee di testi antichi.

Con l'intento di colmare questa lacuna, è nato nel 2018 il progetto *Vertepe*<sup>6</sup> che costituirà il primo *corpus* bilingue della traduzione antica, dedicato in particolare ai testi prosastici greci tradotti in latino in epoca tardoantica<sup>7</sup>.

La banca dati comprenderà ipotesti greci e ipertesti latini, annotati in XML-TEI, di libera e gratuita consultazione su licenza *Creative Commons*, e offrirà come risorse digitali un canone completo e aggiornato con l'indicazione dell'edizione di riferimento delle opere presenti nel *corpus* e un programma di riconoscimento automatico delle corrispondenze greco-latine.

### *Un nuovo metodo di analisi della traduzione antica*

Il progetto *Vertepe* propone un nuovo metodo di analisi della traduzione antica che si avvale da un lato delle risorse e degli strumenti delle *Digital Humanities*, da cui trae ispirazione, e dal-

<sup>5</sup> Numerosi e in continuo aumento sono i *corpora* nel campo degli Studi di traduzione, disciplina che fin dalle sue recenti origini si è servita di metodologie e tecniche non solo della linguistica, ma anche dell'informatica umanistica. Sui *corpora* paralleli di moderne traduzioni si vedano tra gli altri Baker 1993 e Zanettin 1998.

<sup>6</sup> Il progetto è stato realizzato presso il Dipartimento di Studi umanistici dell'Università del Piemonte Orientale e finanziato dalla Fondazione CRT in seguito ad un bando competitivo finalizzato a ricerche nell'ambito della macro-area umanistica. Il sito in costruzione è consultabile al *link* <http://vertere.herokuapp.com>.

<sup>7</sup> Come è noto in questo periodo il greco diventa meno compreso nella parte occidentale dell'impero e quindi si devono elaborare traduzioni anche per le classi più colte. Cfr. Chiesa 2004.

l'altro dei modelli teoretici degli Studi di traduzione<sup>8</sup>. In particolare, il proficuo dialogo tra queste due discipline offre una nuova prospettiva di studio delle traduzioni antiche: il testo digitale non solo permette la precisa classificazione e individuazione dei fenomeni definiti dagli Studi di traduzione ricercabili sia nel singolo testo sia nell'intero *corpus*, ma ne garantisce anche la piena fruizione attraverso una visualizzazione a schermo, libera dai vincoli del supporto cartaceo, che affiancando l'originale alla traduzione ne mette in luce la corrispondenza.

La traduzione dal greco al latino, che può realizzarsi *uerbum de uerbo* o *ad sensum* con riadattamenti che spesso conducono ad una trasformazione radicale del testo di partenza<sup>9</sup>, è stata oggetto di numerosi studi di natura storico-culturale<sup>10</sup> o linguistico-letteraria<sup>11</sup>, che per differente approccio non integrano le strategie individuate dagli Studi di traduzione principalmente volti alla moderna traduttologia<sup>12</sup>.

L'analisi del fenomeno della traduzione negli ultimi decenni, a partire dal fondamentale studio di Genette<sup>13</sup>, ha avuto una rinascita e si devono agli Studi di traduzione la definizione di una terminologia specifica<sup>14</sup> e il merito di aver individuato e classificato delle strategie universali, che possono utilmente applicarsi ai testi antichi, suddivise in metodi diretti (prestiti, calchi, traduzione

<sup>8</sup> La disciplina degli Studi di traduzione, più nota secondo la terminologia anglosassone come *Translation Studies*, affonda le radici nella linguistica teorica degli anni cinquanta del Novecento, nel solco della quale si inseriscono i primi fondamentali contributi di Vinay-Dalbert 1995 [1958], Nida 1964 e Catford 1965.

<sup>9</sup> A riguardo si veda Chiesa 1987.

<sup>10</sup> Si veda in particolare Berschin 2001, pp. 1023-1026.

<sup>11</sup> Fondamentale Bettini 2012.

<sup>12</sup> Costituisce un'isolata eccezione il contributo di Muraru 2011 dedicato a Rufino traduttore di Origene.

<sup>13</sup> Cfr. Genette 1997 [1982].

<sup>14</sup> Nel presente contributo tale terminologia è adottata per quanto riguarda l'originale definito 'ipotesto' e la traduzione definita 'ipertesto'. I due termini non vanno dunque intesi nell'accezione propria delle *Digital Humanities* per la quale si rinvia a Orlandi 2010, p. 25.

letterale) e obliqui (trasposizione, modulazione, equivalenza, adattamento), in strategie sintattico-grammaticali, semantiche e globali-pragmatiche<sup>15</sup>.

Inoltre, solitamente la prospettiva di studio della traduzione antica si concentra su un singolo testo<sup>16</sup> o autore<sup>17</sup>, mentre diversa è la prospettiva dello studio del fenomeno all'interno di un *corpus* con un approccio innovativo basato sulla selezione ed elaborazione di specifiche codifiche XML-TEI che ne descrivano le strategie e ne permettano una disposizione interlineare.

*'Tradurre' in risorsa digitale la traduzione antica*

Una volta individuato l'elenco completo degli ipotesti greci e degli ipertesti latini e redatto un canone con indicazione dell'edizione critica di riferimento, basato rispettivamente sul canone *TLG*<sup>18</sup> con qualche aggiornamento e sul canone della biblioteca digitale *digilibLT*<sup>19</sup>, si è provveduto alla digitalizzazione dei testi greci<sup>20</sup>. Si noti che il progetto *Vertepe*, concepito per la valorizzazione di banche dati esistenti, attinge alle risorse della biblioteca digitale *digilibLT* che offre i testi prosastici della tarda latinità, approssimativamente dalla fine del II sec. d.C. al VII sec. d.C., annotati in XML-TEI, liberamente consultabili e resi disponibili al *download*.

<sup>15</sup> A riguardo si vedano Vinay-Dalbernet 1995 [1958], Malone 1988 e Chesterman 1997.

<sup>16</sup> Tra gli altri si veda, ad esempio, Sconocchia 1984 sul *De plantis*, traduzione latina tardoantica presente nel nostro *corpus*.

<sup>17</sup> A titolo esemplificativo si veda Mazzini 1985 su Ippocrate latino, incluso nel nostro *corpus*.

<sup>18</sup> Cfr. Berkowitz-Squitier 1990.

<sup>19</sup> Cfr. [http://www.digiliblt.uniupo.it/upload/docs/canone\\_con\\_edizioni\\_di\\_riferimento\\_ita.pdf](http://www.digiliblt.uniupo.it/upload/docs/canone_con_edizioni_di_riferimento_ita.pdf). Sul progetto si vedano Tabacco 2014, Tabacco 2016, Lana 2012b, Lana 2012c e Borgna-Musso 2017.

<sup>20</sup> Sono stati ripresi dalla biblioteca digitale *digilibLT* i *files XML* degli ipertesti latini con annotazione strutturale, a cui si è integrata l'annotazione semantica e la codifica di allineamento testuale.

Passaggi obbligati per l'elaborazione digitale sono l'acquisizione delle immagini attraverso la scansione del testo, la conversione delle immagini in un testo UNICODE salvato in formato RTF attraverso il programma OCR *Anagnostis*, la doppia lettura integrale del testo finalizzata alla correzione degli errori di OCR e di eventuali errori di stampa, e infine l'annotazione XML-TEI del testo su vari livelli secondo le specifiche scelte dal progetto.

Per favorire la consultazione *online* delle opere, si è scelto di dotare tutti i testi della codifica di base della struttura formale e dei fenomeni filologici adottata dalla biblioteca digitale *digilibLT*<sup>21</sup>, da cui provengono gli ipertesti latini, integrando ad essa una annotazione semantica descrittiva delle strategie antiche di traduzione e una annotazione di allineamento testuale.

#### *Un'annotazione semantica XML-TEI della traduzione antica*

L'annotazione semantica<sup>22</sup> risponde perfettamente all'esigenza di definire un modello di analisi della traduzione antica, focalizzando l'attenzione sulle strategie di traduzione in un aperto dialogo tra ipotesto greco e ipertesto latino.

L'intento classificatorio degli Studi di traduzione ben si coniuga con questi schemi di codifica: la classificazione più completa delle strategie di traduzione si deve a Chesterman<sup>23</sup> che individua – riprendendo e integrando studi precedenti<sup>24</sup> – dieci strategie sintattico-grammaticali, dieci strategie semantiche e dieci strategie globali-pragmatiche. Si è scelta tale classificazione tripartita come punto di partenza per l'analisi e conseguentemente l'annotazione semantica della traduzione antica.

<sup>21</sup> Cfr. [http://www.digiliblt.uniupo.it/upload/docs/note\\_di\\_trascrizione\\_e\\_di\\_codifica.pdf](http://www.digiliblt.uniupo.it/upload/docs/note_di_trascrizione_e_di_codifica.pdf). A riguardo si veda anche Musso 2014.

<sup>22</sup> Riguardo agli elementi generali per annotazioni interpretative si veda Burnard-Sperberg-McQueen 2005 [2002], pp. 157-162; cfr. anche Orlandi 2010, pp. 136-137.

<sup>23</sup> Si veda il già citato Chesterman 1997.

<sup>24</sup> Cfr. i già citati Vinay-Dalbert 1995 [1958] e Malone 1988.

In un contesto bilingue l'analisi di ipotesto e ipertesto non può limitarsi ad un confronto per parola dal momento che anche le omissioni e le perifrasi giocano un ruolo di rilievo: l'annotazione semantica ha il vantaggio di non essere legata ad un termine specifico e, pertanto, di permettere l'individuazione e il collegamento tra *loci* secondo criteri precisi, dichiarati e ben identificabili, rendendo esplicita e formalizzata l'interpretazione del passo annotato e ampliando la prospettiva di analisi ad una visione d'insieme.

Si è scelto di adoperare una codifica XML-TEI, specifico linguaggio di *markup* per la digitalizzazione dei testi<sup>25</sup>. Tale codifica dunque deve rispettare le norme della sintassi XML, avere coerenza interna ed essere metodologicamente accettabile; a tali vincoli si deve aggiungere la trasparenza di intenti: è infatti determinante esplicitare, nella parte iniziale del *file* chiamata TEI-Header, non solo la modalità di marcatura utilizzata nella codifica dei testi, ma anche la finalità che ha determinato le annotazioni scelte.

Inoltre, l'enorme versatilità e flessibilità del linguaggio XML-TEI, i cui punti di forza sono l'accessibilità, la permanenza nel tempo e la ricercabilità, garantiti dalla standardizzazione della codifica e dall'indipendenza sia da *software* sia da *hardware* specifici, permette non solo di ampliare il ventaglio di codifiche scelte anche in un secondo momento, ma soprattutto consente la sovrapposizione di più codifiche per la medesima porzione testuale a delineare un'interpretazione dettagliata ed esaustiva del passo.

Se dunque il linguaggio XML-TEI è senza dubbio lo strumento ideale per l'analisi del fenomeno della traduzione, è a questo punto spontaneo chiedersi: le strategie di traduzione antiche sono

<sup>25</sup> A sviluppare una codifica *standard* per la rappresentazione dei testi in formato digitale è stato il consorzio di istituzioni internazionali di ambito linguistico-letterario, la TEI, acronimo di *Text Encoding Initiative*. La TEI ricorre al metalinguaggio XML, acronimo di *eXtensible Markup Language*, basato su un meccanismo sintattico che consente di definire e controllare il significato degli elementi contenuti in un testo.

abbastanza rigide da poter essere annotate in una codifica XML-TEI?

Definire l'annotazione semantica della traduzione antica significa convertire le strategie di traduzione in marcatori nel rispetto della tripartizione della struttura classificatoria prevista da Chesterman. In dettaglio, le 'strategie di traduzione' si suddividono in tre macro-gruppi 'strategie sintattico-grammaticali', 'strategie semantiche' e 'strategie globali-pragmatiche'; ciascun macro-gruppo, inoltre, è potenzialmente composto da dieci sotto-categorie interpretative correlate tra loro per tipologia.

Nel frontespizio elettronico, il TEI-header<sup>26</sup>, compaiono le caratteristiche precipe della codifica elaborata: il marcatore <Grp>, *group*, che è utilizzato per unire elementi che hanno qualcosa in comune<sup>27</sup>, è qui utilizzato per definire la macro-categoria delle 'strategie di traduzione'; il marcatore <interpGrp>, *interpretation group*, che viene solitamente usato per unire una serie di interpretazioni correlate, è qui adottato al fine di riproporre nella codifica la suddivisione di Chesterman in tre tipologie tematiche definite attraverso l'attributo @type<sup>28</sup>; infine, dal momento che l'individuazione delle strategie di traduzione nei testi può a tutti gli effetti definirsi un'attività interpretativa, si è scelto di utilizzare il marcatore <interp>, *interpretation*, e di definire le trenta sotto-categorie interpretative specificate attraverso l'attributo @xml:id.

Nel frontespizio elettronico, il TEI-header, si è scelto di inserire all'interno del <fileDesc> dell'ipertesto latino il riferimento esplicito a titolo e edizione di riferimento dell'ipotesto greco e, viceversa, all'interno del <fileDesc> dell'ipotesto greco il riferimento all'ipertesto latino, utilizzando il marcatore <note><sup>29</sup>; inoltre, si è ritenuto opportuno rendere esplicita la lingua dell'opera nel <profileDesc> inserendo per l'ipotesto il marcatore

<sup>26</sup> Sul frontespizio elettronico si veda Burnard-Sperberg-McQueen 2005 [2002], pp. 175-189.

<sup>27</sup> Cfr. TEI *Guidelines* 2019, p. xviii.

<sup>28</sup> Cfr. ivi, § 17.3 *Spans and Interpretations*, pp. 592-596.

<sup>29</sup> Cfr. ivi, § 2.2.6 *The Notes Statement*, p. 35.

<langUsage> e specificando <language ident="gr"> nell'ipotesto e <language ident="lat"> nell'ipertesto<sup>30</sup>.

Si è scelto il marcatore di divisione interna di paragrafo <seg>, *arbitrary segment*, utile per il *mark-up* di porzioni testuali linguisticamente significative e vantaggioso in quanto al suo interno può contenere ulteriori marcatori <seg> per ogni livello di analisi ritenuto appropriato<sup>31</sup>.

Come lingua per la denominazione dei marcatori è stato scelto il latino, ad esempio *'syntaxis grammaticaeque artes'* per *'strategie sintattico-grammaticali'*, inclusivo dei termini traslitterati dal greco, in coerenza con gli obiettivi progettuali, per evitare di impiegare marcatori descritti in lingua inglese abitualmente impiegati in differenti contesti<sup>32</sup>.

La struttura dell'annotazione semantica è stata dunque concepita e definita nel TEI-header come segue:

```
<interp xml:id="synt-gramm">syntaxis grammaticaeque artes</interp>
<interp xml:id="uerb-expl">uerborum explicationis artes</interp>
<interp xml:id="pragm">pragmaticae artes</interp>
```

specificando la suddivisione di ciascun sottogruppo nelle sue sotto-categorie come segue:

```
<interpGrp type="uerb-expl">
  <interp xml:id="synon">synonimia</interp>
  <interp xml:id="auton">autonomia</interp>
  <interp xml:id="hyperon">hyperonimia</interp>
  <interp xml:id="cogit-comm">cogitationis commutatio</interp>
  ...
</interpGrp>
```

Come si può dunque notare, i diversi gruppi di elementi <interp> sono qui raggruppati per tipologia in elementi <interGrp>.

<sup>30</sup> Cfr. *ivi*, § 2.4.2 *Language Usage*, pp. 57-58.

<sup>31</sup> Per l'elemento <seg> cfr. anche *TEI Guidelines 2019*, p. 1591.

<sup>32</sup> In particolare, per l'annotazione di dizionari le *TEI Guidelines* prevedono come attributo @type di <lbl> e di <usg> gli elementi *syn* e *hyper*.



Una volta definita l'annotazione all'interno del TEI-header, i marcatori possono essere utilizzati all'interno del testo, come ad esempio:

```
<seg ana="#synon"> ... </seg>
```

senza la necessità di ripetere in tale sede il sotto-gruppo di appartenenza definito già nel TEI-header. L'attributo @ana contiene un puntatore all'elemento <interp> dove si specifica la strategia di traduzione.

Si noti però che dalle trenta sotto-categorie individuate da Chesterman bisogna escludere le voci 'altri cambiamenti semantici', 'cambiamenti di riscrittura' e 'altri cambiamenti pragmatici' in quanto troppo generiche per prestarsi ad una codifica, e tra le strategie globali-pragmatiche 'filtro culturale', 'cambiamento nella relazione tra testo, autore e lettore' e 'cambiamento di visibilità del traduttore' difficilmente applicabili ai testi antichi e senza riscontro ad una prima ricognizione, riservandosi di introdurle successivamente se presenti.

Le strategie sintattico-grammaticali annotate sono:

traduzione letterale	<i>ad uerbum</i>	<seg ana="#ad-uerb">...</seg>
prestito/calco	<i>mutuatio</i>	<seg ana="#mut">...</seg>
trasposizione	<i>uerborum transgressio</i>	<seg ana="#uerb-trans">...</seg>
cambiamento di unità	<i>unitatis commutatio</i>	<seg ana="#unit-comm">...</seg>
cambiamento di struttura della frase	<i>phrasis structurae commutatio</i>	<seg ana="#phr-struct-comm">...</seg>
cambiamento di struttura del periodo	<i>uerborum ambitus structurae commutatio</i>	<seg ana="#uerb-amb-comm">...</seg>
cambiamento di coesione	<i>coniunctionis commutatio</i>	<seg ana="#con-comm">...</seg>
cambiamento di livello	<i>gradus commutatio</i>	<seg ana="#grad-comm">...</seg>

Una proposta di annotazione semantica

cambiamento di schema retorico	<i>rhetorici schematis commutatio</i>	<seg ana="#rh-sch-comm">...</seg>
--------------------------------	---	-----------------------------------

Le strategie semantiche annotate sono le seguenti:

sinonimia	<i>synonimia</i>	<seg ana="#synon">...</seg>
autonomia	<i>autonomia</i>	<seg ana="#auton">...</seg>
iponimia	<i>hyperonimia</i>	<seg ana="#hyperon">...</seg>
conversi	<i>contraria</i>	<seg ana="#contr">...</seg>
cambiamento di astrazione	<i>cogitationis commutatio</i>	<seg ana="#cogit-comm">...</seg>
cambiamento di distribuzione	<i>distributionis commutatio</i>	<seg ana="#distrib-comm">...</seg>
cambiamento d' enfasi	<i>emphaseos commutatio</i>	<seg ana="#emph-comm">...</seg>
parafrasi	<i>paraphrasis</i>	<seg ana="#paraphr">...</seg>
cambiamento di tropo	<i>figurae commutatio</i>	<seg ana="#fig-comm">...</seg>

Infine, le strategie globali-pragmatiche con riscontro nelle traduzioni antiche sono:

cambiamento di esplicitazione	<i>explicitationis commutatio</i>	<seg ana="#expl-comm">...</seg>
cambiamento di informazione	<i>informationis commutatio</i>	<seg ana="#inform-comm">...</seg>
cambiamento illocutorio	<i>appellationis commutatio</i>	<seg ana="#appel-comm">...</seg>
cambiamento della disposizione logica	<i>logicae locationis commutatio</i>	<seg ana="#log-loc-comm">...</seg>
traduzione parziale	<i>conuersio ex aliqua parte</i>	<seg ana="#conuer-part">...</seg>

Può verificarsi che nell'ipertesto alcune parole non vengano tradotte, ma mantenute così come si trovano nell'ipotesto in caratteri greci o traslitterate in caratteri latini: tale scelta è rilevante ai fini della valutazione complessiva della traduzione.

Pertanto, si è deciso di annotare nell'ipertesto latino i termini greci: a) in caratteri greci con il marcatore `<foreign xml:lang="grc-Grek">...</foreign>`; b) traslitterati in caratteri latini con il marcatore `<foreign xml:lang="grc-Latn">...</foreign>`; nel caso in cui il termine costituisca una strategia di traduzione si procede, invece, aggiungendo allo specifico marcatore `<seg>` rispettivamente l'attributo `@xml:lang="grc-Grek"` e l'attributo `@xml:lang="grc-Latn"`<sup>33</sup>.

### *Un testo antico come modello di codifica*

Sebbene le strategie di traduzione annotate siano ricercabili attraverso una maschera di ricerca avanzata, basilare all'avvio di uno studio critico di una traduzione antica è il confronto *uerbatim* di ipotesto e ipertesto. Pertanto, fondamentale è un affiancamento di ipotesto e ipertesto che metta in evidenza con immediatezza e precisione le corrispondenze testuali greco-latine. Quale tipologia di visualizzazione si presta meglio a tale scopo? Parallela o interlineare?

A consentire di ragionare su tale quesito è un testo del *corpus*: l'*Ars grammatica* greco-latina di Dositeo<sup>34</sup>. Si tratta dunque di un testo molto particolare in quanto Dositeo è autore sia del testo greco sia del testo latino, una traduzione sincronica che la tradizione manoscritta ci trasmette in forma interlineare<sup>35</sup>. L'intento didattico dell'opera, costituita da ipotesto e ipertesto del medesimo autore, gioca a favore dell'aderenza dell'ipertesto all'ipotesto, non solo a livello di resa lessicale ma anche per quanto possibile a livello di *ordo uerborum*. Per tali ragioni, si è scelta questa opera come testo modello per l'elaborazione di un glossario di corri-

<sup>33</sup> Cfr. TEI *Guidelines* 2019, § vi.1 *Language Identification*, pp. xlv-xlvi.

<sup>34</sup> Cfr. Bonnet 2005.

<sup>35</sup> Sebbene non sia dato sapere se la resa interlineare rispecchi la volontà autoriale o sia un espediente introdotto dalla tradizione manoscritta, è indubbio che l'introduzione di tale scelta giochi a favore dell'intento didattico dell'opera concepita per l'apprendimento della lingua latina.

spondenze greco-latine propedeutico allo sviluppo di una codifica di allineamento testuale.

Dunque, in presenza di condizioni ottimali di aderenza testuale, è sicuramente condivisibile e preferibile la scelta della tradizione manoscritta di una disposizione interlineare che conferisce alla resa visiva un immediato confronto testuale – tendenzialmente rigo per rigo – delle corrispondenze greco-latine.

Si è dunque elaborata una codifica di allineamento testuale per la disposizione interlineare: tale annotazione, infatti, non è stata sviluppata nello specifico dalla TEI, che definisce invece i criteri da utilizzare per la visualizzazione testuale parallela<sup>36</sup>, e dunque è stata qui implementata attingendo ai criteri di codifica CES, acronimo di *Corpus Encoding Standard*, e all'uso di marcatori specifici a delimitare i confini della frase.

#### *La sfida dell'automatizzazione*

La visualizzazione interlineare, pur essendo preferibile per i testi antichi, non è comunque pienamente soddisfacente laddove vi siano rilevanti differenze di resa testuale nelle due lingue. Si possono mantenere le corrispondenze greco-latine nella visualizzazione testuale attraverso lo sviluppo di un programma specifico? Può la resa digitale così superare i limiti del supporto cartaceo?

A tale scopo si è sviluppato un programma che automatizza il riconoscimento delle corrispondenze testuali greco-latine per parola tra ipotesto e ipertesto attraverso concordanze parallele, ossia sulla base di un repertorio preimpostato di corrispondenze tra parole ed espressioni nelle due lingue.

Semplificando, le corrispondenze tratte dall'elaborazione dei dati del glossario dell'*Ars grammatica* di Dositeo sono oggetto di una mappatura nella flessione non con semplice metodo statistico ma con ancore linguistiche, e unitamente si procede all'assegnazione di un codice di identificazione ad ogni livello sintattico.

<sup>36</sup> Cfr. TEI *Guidelines* 2019, § 16 *Linking, Segmentation, and Alignment*, in particolare 16.5.2 *Alignment of Parallel Texts*, pp. 558-560.

Tale programma può definirsi ‘autodidatta’ in quanto all’inserimento dei dati bilingui è seguita una fase di auto-addestramento durante la quale il programma ha imparato a riconoscere autonomamente altre corrispondenze sulla base di analogie di contesto e sintassi. Per la prima fase di ‘addestramento’ del programma si è scelta infatti una rosa di testi appartenenti a generi letterari differenti in modo da coprire un ampio ventaglio lessicale.

Maggiore è il numero di testi riconosciuti, maggiore è il grado di esattezza del riconoscimento automatico: determinante è dunque che il *corpus* sia integralmente digitalizzato e attualmente si sta procedendo in questa direzione.

### *Conclusioni*

L’annotazione semantica qui presentata intende essere una proposta aperta, da cui il titolo del presente contributo, ispirata ai principi della collaborazione e della condivisione che governano le *Digital Humanities*.

Dal proficuo dialogo tra le risorse digitali *online*, accomunate dal medesimo linguaggio XML-TEI e dalla natura interdisciplinare, nascono nuove prospettive di ricerca. Il progetto *Vertepe* affonda le radici nell’assidua ‘frequentazione’ della già citata biblioteca digitale *digilibLT* e della banca dati *TBL, Textual Bilingualism in Latin*<sup>37</sup> che indaga i fenomeni di contatto linguistico tra greco e latino nei testi letterari della tarda latinità, e guarda al futuro non solo cercando di dialogare con altri progetti, ma anche proponendo nuovi approcci alla ricerca, tra i quali il progetto dal titolo *Raccontare il mito nell’antichità: una proposta di annotazione XML-TEI del riuso testuale*. Si tratta di un progetto recentemente finanziato dalla Fondazione CRT, concepito ed elaborato proprio a partire dall’analisi del *corpus* di *Vertepe*, con il duplice scopo di indagare il fenomeno del ‘riuso testuale’, da

<sup>37</sup> Cfr. <http://tbl.uniupo.it/index.php#>. Sul progetto si vedano Fedriani-Napoli-Rosso 2019 e Fedriani-Napoli-Rosso (in corso di pubblicazione).

intendersi come la ripresa di porzioni testuali identiche o simili in testi differenti, nella mitografia di consultazione greco/latina, e di elaborare un metodo innovativo di analisi basato sullo sviluppo di un programma di automatizzazione dell'individuazione dei *loci similes* mitografici e sulla definizione di una codifica XML-TEI che definisca il livello di richiamo.

Si potrebbe forse guardare all'odierno panorama digitale come ad un tessuto, costituito da molteplici fili che se presi singolarmente mantengono la propria specificità, ma solo se intrecciati tra loro formano la trama.

**Abstract.**

This paper provides a new method of analysis of ancient translations which draws on key concepts from *Digital Humanities* and *Translation studies*. The research focuses on a group of translations from Greek into Latin which have been selected as especially representative of this phenomenon in Late antiquity. The purpose of this paper is to define a specific XML-TEI annotation for ancient translation. The annotation shows and classifies the main typologies of translation strategies present in the corpus. The paper also illustrates IT technologies that allow an interlinear visualisation of hypotext and hypertext of Greek and Latin passages as it is shown in the manuscripts of *Ars grammatica* by Dositheus. The new annotation and the corpus-based approach combine quantitative and qualitative perspectives, offering a new approach to the study of ancient translation techniques.

**Keywords.**

Greek-Latin translation, XML-TEI annotation, translation strategies.

Nadia Rosso

Università del Piemonte Orientale

nadia.rosso@uniupo.it

## BIBLIOGRAFIA

- Baker 1993: M. Baker, *Corpus Linguistics and Translation Studies. Implications and Applications*, in M. Baker-G. Francis-E. Tognini-Bonelli (ed. by), *Text and Technology. In Honour of John Sinclair*, John Benjamins Publishing, Philadelphia-Amsterdam, pp. 233-252.
- Berkowitz-Squitier 1990: L. Berkowitz-K.A. Squitier, *Thesaurus linguae Graecae Canon of Greek Authors and Works*, Oxford University Press, New York-Oxford.
- Berschin 2001: W. Berschin, *Traduzioni dal greco in latino (secoli IV-XIV)*, in S. Settis (a cura di), *I Greci. Storia, cultura, arte, società*, vol. III, *I Greci oltre la Grecia*, Einaudi, Torino, pp. 1023-1033.
- Bettini 2012: M. Bettini, *Vertere. Un'antropologia della traduzione nella cultura antica*, Einaudi, Torino.
- Bonnet 2005: *Dositheus. Ars grammatica. Dosithee, Grammaire latine*, texte établi, trad. et commenté par G. Bonnet, Les Belles Lettres, Paris.
- Borgna-Musso 2017: A. Borgna-S. Musso, *Le sfide di una biblioteca digitale del latino tardoantico. Struttura, canone e questioni aperte di codifica e visualizzazione*, in P. Mastandrea (a cura di), *Strumenti digitali e collaborativi per le Scienze dell'Antichità*, Edizioni Ca' Foscari, Venezia.
- Burnard-Sperberg-McQueen 2005: L. Burnard-C.M. Sperberg-McQueen, *Il manuale TEI Lite. Introduzione alla codifica elettronica dei testi letterari*, [2002], trad. it. F. Ciotti, Edizioni Sylvestre Bonnard, Milano [[http://www.tei-c.org/Vault/P4/Lite/teiu5\\_it.html](http://www.tei-c.org/Vault/P4/Lite/teiu5_it.html) (2019-12-12)].
- Catford 1965: J.C. Catford, *A Linguistic Theory of Translation*, Oxford University Press, Oxford.
- Chesterman 1997: A. Chesterman, *Memes of Translation. The Spread of Ideas in Translation Theory*, Benjamin Publishing, Amsterdam-Philadelphia.
- Chiesa 1987: P. Chiesa, *Ad uerbum o ad sensum? Modelli e coscienza metodologica della traduzione tra tarda antichità e alto medioevo*, «Medioevo e Rinascimento» 1, pp. 1-51.
- Chiesa 2004: P. Chiesa, *Le traduzioni in latino di testi greci*, in M. Capaldo et al. (a cura di), *Lo spazio letterario del medioevo*, vol. III, *Le culture circostanti, I, La cultura bizantina*, Salerno Editrice, Roma, pp. 491-518.
- Fedriani-Napoli-Rosso 2019: C. Fedriani-M. Napoli-N. Rosso, *Forms and functions of Greek words in Late Latin literary texts: a corpus-based approach*, in N.

*Una proposta di annotazione semantica*

Holmes-M. Ottink-J. Schrickx (ed. by), *Lemmata Linguistica Latina. Words and Sounds*, vol. I, De Gruyter, Berlin, pp. 294-311.

Fedriani-Napoli-Rosso (in corso di pubblicazione): C. Fedriani-M. Napoli-N. Rosso, *Un nuovo corpus per lo studio del bilinguismo testuale in latino tardo. Aspetti metodologici e primi risultati*, in *Latin vulgaire latin tardif XII, Actes du XIIIe Colloque international sur le latin vulgaire et tardif, Uppsala, August 2016*.

Genette 1997: G. Genette, *Palimpsesti. La letteratura al secondo grado*, [1982], trad. it. R. Novità, Einaudi, Torino.

Lana 2012a: M. Lana, *Biblioteche digitali. Un'introduzione*, Bononia University Press, Bologna.

Lana 2012b: M. Lana, *Metodologie e problematiche per una biblioteca digitale. Il caso di digilibLT*, *DigiItalia*, «Rivista del digitale nei beni culturali» 7, pp. 40-64.

Lana 2012c: M. Lana, *Da una digital library del latino tardo ad un corpus globale*, in F. Ciotti-G. Crupi (a cura di), *Dall'informatica umanistica alle culture digitali, Atti del convegno di studi in memoria di Giuseppe Gigliozzi, Roma, 27-28 ottobre 2011*, Digilab, Roma, pp. 134-150.

Malone 1988: J.L. Malone, *The Science of Linguistics in the Art of Translation*, State University of New York Press, Albany.

Mazzini 1985: I. Mazzini, *Ippocrate latino dei secoli V-VI. Tecnica di traduzione*, in I. Mazzini-F. Fusco (a cura di), *I testi di medicina latini antichi. Problemi filologici e storici, Atti del I Convegno Internazionale, Macerata-S. Severino M., 26-28 aprile 1984*, G. Bretschneider, Roma, pp. 383-387.

Muraru 2011: A. Muraru, *Strategies of translation in late antiquity: Rufinus and the bilingual readers of Origen's Homilia in exodum 9*, «Adamantius» 17, pp. 297-302.

Musso 2014: S. Musso, *La marcatura di testi latini tardoantichi. Un compromesso ragionevole*, «Aevum Antiquum» 11, pp. 151-187.

Nida 1964: E. Nida, *Toward a Science of Translating: with Special Reference to Principles and Procedures Involved in Bible Translating*, Brill, Leiden.

Orlandi 2010: T. Orlandi, *Informatica testuale. Teoria e prassi*, Laterza, Roma-Bari.

Sconocchia 1984: S. Sconocchia, *Problemi di traduzione del testo greco del De plantis duodecim signis et septem planetis subiectis attribuito a Tessalo di Tralle. I rapporti tra la traduzione latina tardoantica e la traduzione latina medioevale*, in G. Sabbah (éd. par), *Textes médicaux latins antiques*, Publications de l'Université, Saint Etienne, vol. V, pp. 125-151.



Nadia Rosso

Tabacco 2014: R. Tabacco, *La tarda antichità latina tra i codici e il web*, «Aevum Antiquum» 11, pp. 3-17.

Tabacco 2016: R. Tabacco, *La prosa latina pagana tardoantica e la biblioteca digitale digilibLT*, in C. Marazzini-L. Maconi (a cura di), *L'italiano elettronico. Vocabolari, corpora, archivi testuali e sonori, Atti del Convegno "Piazza delle Lingue", VIII Edizione, Firenze, 6-7 novembre 2014*, Edizioni dell'Accademia della Crusca, Firenze, pp. 125-141.

Vinay-Darbelnet 1995: J.P. Vinay-J. Darbelnet, *Contrastive Stylistics of French and English. A Methodology for Translation*, [1958], trad. ingl. J.C. Sager-M.J. Hamel, John Benjamins Publishing, Amsterdam-Philadelphia.

Zanettin 1998: F. Zanettin, *Bilingual Comparable Corpora and the Training of Translators*, in S. Laviosa (ed. by), *L'approche basée sur le corpus / The Corpus-Based Approach*, «META» XLIII, 4, pp. 616-651.

TEI Guidelines 2019: *Guidelines for Electronic Text Encoding and Interchange* [<https://tei-c.org/guidelines/p5/> (2019-12-12)].