

Sviluppo di un Corpus digitale del Lineare A

Relazione finale per il Seminario di Cultura Digitale

Tommaso Petrolito

LM in Informatica Umanistica, matricola 441099

22-07-2015

1 Introduzione

Questa relazione descrive il corpus digitale del Lineare A (o del minoico, che è la lingua sconosciuta che sarebbe trascritta con tale sistema di scrittura) e l'approccio che è stato seguito per svilupparlo. La speranza è di fornire ai curiosi e agli accademici una utile risorsa di studio e d'indagine.

In primo luogo si procederà a un'introduzione all'argomento dal punto di vista storico, descrivendo il Lineare A e il minoico, per chiarire perché si dovrebbe sviluppare un corpus digitale delle trascrizioni in Lineare A esistenti.

In secondo luogo saranno elencate le risorse esistenti relative al Lineare A: i "documenti" in Lineare A (sigilli, statuette, vasellame etc.), i sistemi di trascrizione e traslitterazione utilizzati tradizionalmente dagli studiosi del Lineare A (codici numerici convenzionali relativi a ciascun simbolo), un vecchio font del Lineare A e il set di caratteri Unicode recentemente (16 Giugno 2014) rilasciato per il Lineare A.

In terzo luogo sarà esposta la scelta relativa al formato dei dati: perché è stato deciso di digitalizzare le risorse del Lineare A; perché è stato deciso di convertire le trascrizioni/traslitterazioni in caratteri Unicode; perché è stato deciso di utilizzare un formato XML ed in particolare la DTD TEI-EpiDoc.

Infine saranno descritti: il processo di sviluppo (dalla raccolta dei dati ai problemi incontrati e alle soluzioni trovate); un nuovo font sviluppato appositamente (allineato con le corrette posizioni nel set di caratteri Unicode) al fine di rendere i dati leggibili anche su sistemi e software non aggiornati.

A chiusura della relazione seguirà una discussione più teorica su due aspetti: da un lato una riflessione sulla necessità della standardizzazione (per quanto concerne il formato e la codifica) e della diffusione della risorsa in una prospettiva di open data, dall'altro una riflessione sul valore di corpora epigrafici, come quello in questione, in termini di patrimonio culturale, anche al di là del contenuto strettamente linguistico (a maggior ragione in questo specifico caso, non essendo stato decifrato il minoico).

Sistema di scrittura	Arco temporale
Geroglifico cretese	2100 – 1700 BC
Lineare A	2500 – 1450 BC
Lineare B	1450 – 1200 BC

Tabella 1: Archi temporali del Geroglifico cretese, il Lineare A e il Lineare B.

2 Il Lineare A e il minoico

Il Lineare A è il sistema di scrittura usato dalla civiltà Minoica [8] dal 2500 al 1450 BC.

La civiltà minoica fiorì sull'isola di Creta nel mare Egeo durante l'Età del Bronzo. Le rovine minoiche sono state ritrovate principalmente a Creta ma anche in altre isole greche e in Grecia continentale, oltre che in Bulgaria, in Turchia e in Israele.

Il Lineare A non è ovviamente più utilizzato e, nonostante siano passate decenni di studi (fu scoperto da Arthur Evans intorno al 1900 [13]), rimane ancora indecifrato.

Tutte le ipotesi avanzate riguardo al Lineare A e al minoico sono basate principalmente sulla comparazione con il molto più conosciuto Lineare B, il famoso sistema di scrittura "figlio" derivato dal Lineare A. Infatti, il Lineare B è stato pienamente decifrato durante gli anni '50 da Michael Ventris¹, si è scoperto che esso era utilizzato per trascrivere un antico dialetto del greco antico parlato dalla civiltà micenea.

L'archeologo Arthur Evans definì il sistema di scrittura "Lineare" poiché esso consisteva di semplici segni dalla forma lineare, linee, incise nella terracotta e nella pietra [34] mentre, nello stesso periodo (come si può osservare nella Tabella 1), il Geroglifico cretese era più pittografico e tridimensionale.

Anche se molti simboli sono condivisi dal Lineare A e dal Lineare B, non è stato possibile individuare parole intellegibili (in alcuna lingua conosciuta), nelle iscrizioni in Lineare A, applicando la segmentazione e i valori fonetici del Lineare B.

Il Lineare A consiste di centinaia di simboli, alcuni dei quali probabilmente con valore sillabico, altri con valore ideografico. Si pensa che molti dei simboli del Lineare A che si trovano anche nel Lineare B (81 in totale) abbiano valore sillabico, mentre gli altri sono considerati presumibilmente logogrammi.

Ci sono stati molti tentativi di decifrazione del Lineare A e del minoico. Si possono suddividere le ipotesi in sei gruppi: una lingua basata sul Greco

¹Si consiglia, in merito, la lettura di questo articolo nella sezione delle notizie relative alla ricerca scientifica del sito web dell'Università di Cambridge: <http://www.cam.ac.uk/research/news/cracking-the-code-the-decipherment-of-linear-b-60-years-on>

Text	𐀓𐀠𐀥
Phonetic Value	<i>KU-NI-SU</i>
Possible Meaning	<i>Knossos</i>

Tabella 2: Esempio di un tentativo di John G. Younger di decifrazione di toponimi.

[25], una lingua Indoeuropea ma appartenente ad un ramo autonomo [27], una lingua anatolica vicina al Luvio [29], una lingua arcaica simile al Fenicio [10], una lingua Indoiranica [15] e una lingua simile all'Etrusco [14].

Esiste anche un interessante tentativo [41] di decifrazione di singole parole, in particolare toponimi, applicando i valori fonetici del Lineare B ai simboli condivisi da Lineare A e Lineare B e seguendo l'ipotesi che i toponimi abbiano avuto una probabilità molto maggiore di sopravvivere come prestiti linguistici nel greco di Micene (scritto in Lineare B); un esempio di questo approccio è mostrato nella Tabella 2.

Nelle prossime sezioni saranno descritti il corpus digitale del Lineare A (perché e come è stato sviluppato) e le risorse già esistenti e disponibili sul Lineare A.

3 Risorse disponibili sul Lineare A

Anche se il Lineare A e il Lineare B sono stati scoperti più di un secolo fa, il Lineare A non è stato ancora decifrato. Nonostante ciò, molti studiosi hanno lavorato sulla raccolta ed organizzazione di tutti i "documenti" e di tutte le informazioni, al fine di studiare e decifrare sia la lingua che il sistema di scrittura.

Probabilmente per il fatto che solamente linguisti storici, filologi e archeologi hanno tentato di raccogliere e organizzare i dati esistenti, ancora tutt'oggi un corpus digitale del Lineare A ricco e ben organizzato non esiste.

In questa sezione sono descritte tutte le risorse esistenti per il Lineare A, siano esse risorse "fisiche" o digitali.

3.1 I "documenti" trascritti in Lineare A

Il Lineare A fu trascritto su una grande varietà di supporti: tavole di pietra per offerte religiose, forcine d'oro e d'argento e vasellame (sia scritto con inchiostri/pigmenti, sia inciso).

I documenti in terracotta consistono di tavolette, tondini, e sigilli (monoforori, bifori e a base piatta).

I tondini hanno a che fare con "conveyance of a commodity, either within the central administration or between the central administration and an

ID	Tipo di supporto
default	tavolette
Wa	noduli
Wb	sigilli
Wc	tondini
Za	vasellame di pietra
Zb	vasellame di terracotta
Zc	iscrizioni colorate
Zd	graffiti
Ze	elementi architettonici
Zf	oggetti di metallo
Zg	oggetti di pietra

Tabella 3: Indicizzazione del tipo di supporto [44].

external party" ("la registrazione del trasferimento di un bene di proprietà entro l'amministrazione centrale o tra l'amministrazione centrale e un terzo") [30, 35]. Il tondino registra transazioni regolamentate dall'amministrazione centrale quando un bene esce dal luogo di transazione, come una moderna ricevuta [19]. I sigilli bifori identificavano i beni portati verso il centro; i sigilli monofori erano apposti su documenti di papiro o pergamena; i sigilli a base piatta (mai iscritti) venivano pressati sullo spago che chiudeva i documenti di papiro e pergamena [46, 35] come mostrato da fotografie [23, 24] delle stampe che sopravvivono sul lato inferiore dei sigilli a base piatta.

Esistono 1.427 documenti in Lineare A contenenti tra 7.362 e 7.396 simboli (in base alle interpretazioni), quantità molto minore di quella disponibile per il Lineare B (più di 4.600 documenti contenenti 57.398 simboli) [45].

3.2 La Raccolta di Iscrizioni in Lineare A di Godart e Olivier

Esiste una raccolta cartacea completa ed organizzata di documenti in Lineare A, la raccolta **GORILA** (Louis **G**odart and Jean-Pierre **O**livier, *Recueil des inscriptions en Linéaire A* [17]).

Godart e Olivier hanno catalogato ed indicizzato i documenti per luogo di origine e tipo di supporto, seguendo l'Indice Raison-Pope [31].

Ad esempio, il documento **AP Za 1** è stato rinvenuto a **AP** = *Apodoulou* e il tipo di supporto è **Za** = *vasellame di pietra* come mostrato nella Tabella 3.

John G. Younger [47] fornisce una mappa dei siti di ritrovamento sull'isola di Creta e una mappa con tutti i siti di ritrovamento greci fuori dall'isola di Creta [48].

Godart e Olivier forniscono anche informazioni riguardo ai luoghi di conservazione (principalmente musei) e alla periodizzazione (per esempio: **EM II** = *Second Early Minoan*, in Italiano: **AM II** = *Antico Minoico Secondo*).

Dal 1976, questa è stata la principale fonte di informazioni e punto di riferimento riguardo i documenti in Lineare A, ponendo le basi per studi più approfonditi. Anche raccolte più recenti, quali il *Corpus transnuméré du linéaire A* [32] (venti anni dopo), fanno sempre riferimento agli specifici volumi e alle specifiche pagine di GORILA che descrivono ogni documento.

3.3 Il sito web di John G. Younger

John G. Younger ha pubblicato un sito web [49] che costituisce, al momento, la migliore risorsa digitale disponibile (esiste anche un altro progetto molto interessante, mai completato, sul sito di Yannis Deliyannis²). Raccoglie la maggior parte delle iscrizioni esistenti (prendendo GORILA come principale fonte di informazioni e punto di riferimento) trascritte come valori fonetici del Lineare B (come nella trascrizione *KU-NI-SU* vista sopra) e come codici GORILA (si veda la Sottosezione 3.4 in merito) nel caso di simboli inclusi solo nel Lineare A.

3.4 Catalogo di simboli GORILA

Sono stati definiti vari sistemi di trascrizione dai diversi studiosi del Lineare A.

Il primo è stato proposto da Raison e Pope [31] e utilizza una stringa composta da uno o due caratteri (*Lm*, *L* o *Lc* in base al tipo di simbolo, rispettivamente metrico, fonetico o composto) seguito da un numero, ad esempio: *L2*.

Questo sistema è stato ampiamente utilizzato da molti studiosi tra i quali David Woodley Packard (presidente del Packard Humanities Institute³ e figlio di David Packard, cofondatore della Hewlett-Packard), Colin Renfrew and Richard Janko [28, 33, 21].

Il secondo, utilizzato nella raccolta GORILA [17] e nel sito web di John G. Younger, consiste di una stringa composta da uno o due caratteri (*AB* se il simbolo è condiviso da Lineare A e Lineare B, *A* se il simbolo è usato solo nel Lineare A) seguito da un numero ed eventualmente da altri caratteri alfabetici (in ragione di *addenda* e *corrigenda* alle versioni precedenti), ad esempio: *AB03*.

Molti studiosi trascrivono i simboli condivisi da Lineare A e Lineare B con l'ipotetica trascrizione fonetica/sillabica. Questa trascrizione sillabica è basata sui valori fonetici corrispondenti in Lineare B. John G. Younger fornisce una tabella di conversione [40] dei vari sistemi di trascrizione descritti:

²Il sito è visitabile all'URL: <http://y.deliyannis.free.fr/linearA/>

³Il cui sito web è visitabile alla pagina: <http://www.packhum.org/>

Unicode	GORILA	Pope&Raison	Sillaba
10600	AB01	L30	DA
10601	AB02	L22	RO
10602	AB03	L2	PA

Tabella 4: Estratto dalla tabella di conversione tra i vari tipi di trascrizione di John G. Younger (GORILA - Pope&Raison - valore fonetico) esteso con la codifica Unicode.

quello di Pope&Raison, quello di GORILA e il suo sistema di trascrizione fonetica/sillabica.

Sviluppando il corpus, sono state prese in considerazione principalmente le trascrizioni fonetiche/sillabiche di John G. Younger, tenendo in considerazione la mappatura sui codici GORILA, poiché il set di caratteri di Unicode per la codifica del Lineare A è ampiamente basato sul catalogo di simboli/codici di GORILA⁴, che è anche il set principale di codifica/trascrizione utilizzato nei tentativi di decifrazione. Un esempio delle diverse trascrizioni per lo stesso simbolo è disponibile nella Tabella 4. Come si può notare, il set Unicode è basato sull'ordine del sistema di trascrizione di GORILA.

3.5 LA.ttf: un font per il Lineare A

Il migliore font per il Lineare A Font disponibile è `LA.ttf`, rilasciato da D.W. Borgdorff⁵ nel 2004.

In questo font, i simboli del Lineare A sono mappati su posizioni Unicode arbitrarie corrispondenti a caratteri latini.

Da un lato ciò consente all'utente di digitare simboli del Lineare A direttamente premendo i pulsanti sulla tastiera; dall'altro, è possibile produrre solo mere traslitterazioni. Il testo digitato internamente resta comunque una serie di caratteri latini e nulla di più.

Va precisato che questo font, dunque, non sarebbe utile per rendere leggibile un corpus del Lineare A codificato in Unicode e non traslitterato tramite i caratteri latini.

⁴Si confrontino, in merito, le specifiche ufficiali della versione 7 dello standard unicode (<http://www.unicode.org/versions/Unicode7.0.0/ch08.pdf>), si noterà l'esplicito riferimento al catalogo GORILA e al suo ruolo di fonte primaria per il set di caratteri del Lineare A.

⁵Reperibile all'URL: <http://www.fontineed.com/author/D.W.%20Borgdorff>

3.6 Il set di caratteri Unicode per il Lineare A

In data 16 Giugno 2014 è stata rilasciata la versione 7.0 dello standard Unicode⁶, aggiungendo 2.834 nuovi caratteri incluso, finalmente, il set di caratteri per il Lineare A.

Il blocco di caratteri per il Lineare A è stato fissato nell'intervallo 10600–1077F e l'ordine segue quasi alla perfezione l'ordine dei codici GORILA come esplicitato dalla documentazione ufficiale⁷ ed evidente sulla Tabella 4.

Il set Unicode copre i segni semplici, quelli a forma di vaso, i segni complessi (composti), i segni complessi includenti forme di vaso, frazioni e composti di frazioni.

Questa risorsa apre, per la prima volta, la possibilità di sviluppare un corpus digitale del Lineare A non consistente in mere traslitterazioni o trascrizioni alternative (quali quella fonetica).

4 Formato del Corpus

Molti studiosi si sono confrontati con i problemi relativi alla *data curation* e hanno considerato varie possibilità.

Fra tutte le possibili soluzioni, è stato scelto di sviluppare un corpus digitale del Lineare A come una raccolta di documenti XML con schema di codifica TEI-EpiDoc.

In questa sezione sono spiegate le ragioni di tale scelta.

4.1 Perché digitale?

Molti studiosi di epigrafia cominciano a digitalizzare corpora epigrafici, in effetti ci sono varie ragioni per cui digitalizzare un corpus epigrafico è decisamente una buona idea.

Un corpus digitale può includere molte rappresentazioni delle iscrizioni [22]:

- foto del documento originale;
- immagini o disegni a mano per avere un'immagine semplificata dell'aspetto del documento;
- trascrizioni diplomatiche;
- modifiche al testo;

⁶Il rilascio fu annunciato sul blog ufficiale di Unicode in questo articolo: <http://blog.unicode.org/2014/06/announcing-unicode-standard-version-70.html>

⁷Consultando la documentazione ufficiale (<http://www.unicode.org/versions/Unicode7.0.0/ch08.pdf>) si legge: "The repertoire of characters in the Unicode encoding of the Linear A script is broadly based on the GORILA catalog by Godart and Olivier"

- traduzioni;
- commento.

Costruire una base di dati sarebbe già abbastanza per ottenere funzionalità molto più ricche di quelle fornite da un corpus cartaceo. La funzionalità più evidente di una base di dati epigrafica è la sua indipendenza da una limitata varietà di indici di ricerca, come sarebbe invece per gli indici analitici di un'edizione cartacea. Si tratterebbe di un *Index Universalis* [18], a differenza degli indici scritti a mano, non vincolato a mantenere solo un numero ristretto di indici di ricerca.

Una funzionalità da tenere in considerazione è l'ovviamente preziosa opportunità di rendere i dati disponibili anche sul web.

4.2 Perché Unicode?

L'elaborazione digitale dei testi deve tenere conto del sistema di scrittura rappresentato nel corpus. Se il corpus consiste di iscrizioni scritte in alfabeto latino, allora il sistema di scrittura delle iscrizioni corrisponderà allo stesso sistema di scrittura delle lingue moderne dell'Europa occidentale usato per i metadati le traduzioni e i commenti al testo. Nel caso in questione, sfortunatamente, si ha a che fare con il Lineare A, dunque c'è bisogno di trovare un modo di rappresentare il testo. Infatti, gli studiosi hanno avuto da obiettare alle basi di dati epigrafiche, sul piano della loro scarsa potenzialità grafica di rappresentare sistemi di scrittura non latini [1].

Ciò ha portato, in alcune basi di dati, all'utilizzo di font non standard, cosa che s'è dimostrata controproducente, compromettendo la compatibilità globale e l'aggiornamento del sistema. Questo approccio è molto invitante poiché, se il corpus dev'essere stampato, prima o poi il font servirà in ogni caso. La soluzione basata sul font presuppone che tutti i software coinvolti nei processi di elaborazione possano riconoscere i marcatori di cambio del font. Sfortunatamente, alcuni Database Management System non consentono cambi di font entro un campo e alcuni formati di esportazione perdono del tutto l'informazione relativa al font.

Qualora i sistemi di scrittura necessari siano tutti supportati, e sarà certamente il caso di ogni sistema di scrittura ancora utilizzato da lingue vive, Unicode costituirebbe certamente un approccio più vantaggioso. Nonostante il minoico non sia una lingua viva, il Lineare A, come già anticipato, è finalmente parte dello standard Unicode⁸, anche se alcuni gruppi di segni convenzionalmente interpretati come numeri non hanno rappresentazione Unicode.

⁸Un elenco dettagliato delle lingue vive, morte o inventate per le quali la codifica Unicode offre il supporto è disponibile alla pagina: <http://www.unicode.org/charts/>

4.3 Perché XML?

Sino a non molto tempo fa, i sistemi di annotazione in questo settore hanno sempre richiesto speciali simboli tipografici, parentesi, puntini, etc.

Alcuni corpora possono utilizzare in particolare i simboli tipografici del sistema di Leida (o "convenzione di Leida") [20, 39, 38] che ha il vantaggio di essere pienamente familiare agli studiosi di epigrafia, cioè coloro che creano e mantengono il corpus. Sfortunatamente, le parentesi speciali, i puntini speciali e altri strumenti tipografici potrebbero non essere supportati dal set di caratteri del sistema.

Alcuni epigrafisti vedono l'annotazione XML come una naturale trasformazione di ciò che loro hanno sempre fatto, con in più i vantaggi offerti dalla standardizzazione entro una vasta comunità di annotatori.

C'è un consenso crescente sul fatto che XML sia il modo migliore di annotare i testi. Un vantaggio chiave dell'utilizzo di XML è la possibilità di scambiare i dati con altri progetti, come anche è utilissimo poter dividere l'informazione su più livelli di annotazione: catalogazione, annotazione, commento, modifiche.

In alcuni casi, potrebbe essere necessario integrare livelli diversi da progetti diversi (ad esempio nel caso in cui ciascuno di questi progetti sia orientato su un diverso livello specifico, per il quale fornisca la migliore qualità), di conseguenza i dati dei vari progetti e del progetto finale, risultato dell'integrazione dai diversi progetti, dovrebbero essere in fatti compatibili.

Se i progetti, basati su annotazione XML, usassero per giunta la stessa DTD, ottenere questo risultato sarebbe relativamente semplice.

Se è vero che anche corpora che conservino i loro testi come semplici documenti Word (o analoghi) con annotazione basata sul sistema di Leida possono condividere i dati tra loro, è da tenere presente che, in tal caso, per ottenere questo risultato, i corpora dei diversi progetti dovranno conformarsi esplicitamente e nei dettagli per quanto riguarda la struttura del testo, il formato dei file, e la codifica dei caratteri. Inoltre, con l'annotazione XML, è possibile definire elementi o entità per rappresentare caratteri non supportati.

Questa caratteristica è particolarmente interessante nel caso in questione, fornendo una soluzione per la rappresentazione dei numeri (i gruppi di segni del Lineare A convenzionalmente interpretati come numeri, a eccezione delle frazioni, non hanno rappresentazione Unicode). Si supponga di voler annotare in XML il gruppo di segni "5", convenzionalmente interpretato come il numero 5. Come specificato nelle DTD TEI, questo potrebbe essere rappresentato come `<g ref="#n5"/>`, dove l'elemento `g` indica un glifo o un carattere non standard e il valore dell'attributo `ref` si riferisce all'elemento `glyph`, contenente informazioni riguardo lo specifico glifo. Un esempio è mostrato nella Figura 1.

```

<glyph xml:id="n5">
  <glyphName>
    Number 5
  </glyphName>
  <mapping type="standardized">
    5
  </mapping>
</glyph>

```

Figura 1: Esempio di dichiarazione XML di un glifo sprovvisto di rappresentazione Unicode.

In alternativa, la DTD potrebbe definire un'entità per rappresentare questo carattere.

In entrambi i casi, il testo XML tiene nota del fatto che è presente un numero 5 in Lineare A e il successivo rendering del testo per una versione da distribuire stampata o a schermo può procedere a sostituire gli elementi `g` o i riferimenti a entità rappresentanti il numero 5 nel testo con il carattere appropriato in un font conosciuto, un'immagine del carattere o anche un numero in un altro sistema di scrittura. Approcci di questo tipo richiedono che ci siano strumenti disponibili per questo tipo di conversioni; un'applicazione o un foglio di stile devono saper come interpretare gli elementi o le entità definite.

Riassumendo, dunque, l'utilizzo dell'XML fornisce due vantaggi fondamentali: in primo luogo, è veramente utile per codificare meta-informazioni anche su più livelli; in secondo luogo, rende possibile la codifica dei caratteri che occorrono nel testo (come mostrato sopra).

4.4 Perché EpiDoc?

Se un gruppo di ricerca decide di annotare in XML, è necessario che venga scelta la DTD o lo schema che si voglia utilizzare. Come in ogni altra disciplina umanistica (non solo nel caso dell'epigrafia nello specifico), la domanda basilare è se utilizzare una DTD generica, come la TEI, o se scrivere una DTD specifica per il progetto in questione. Alcuni progetti richiedono DTD molto specifiche per i tipi di iscrizioni con i quali si ha a che fare, invece altri progetti preferiscono contare su DTD esistenti molto diffuse e supportate.

Mahoney ha analizzato approfonditamente tutte le questioni e le problematiche legate alla digitalizzazione [22] tenendo conto di tutti i vantaggi e gli svantaggi dei diversi approcci; la sua conclusione è che la cosa migliore sia utilizzare **EpiDoc**⁹ uno strumento di codifica XML che potrebbe essere uti-

⁹Una presentazione del progetto è consultabile alla pagina: <http://www.tei-c.org/Activities/Projects/ep01.xml>

lizzato anche per scrivere documenti strutturati conformemente allo standard TEI¹⁰.

La DTD EpiDoc è la DTD TEI, con un limitato numero di personalizzazioni specifiche orientate all'epigrafia ed è stata realizzata utilizzando i meccanismi standard TEI. Piuttosto che scrivere una DTD per l'epigrafia da zero, il gruppo EpiDoc ha deciso di utilizzare TEI perché TEI aveva già risolto molte delle sfide tassonomiche e semantiche affrontate dagli epigrafisti; inoltre, la comunità di utenti TEI può fornire un vasto repertorio di esempi di buone pratiche e suggerimenti. Infine, il repertorio di strumenti costruito intorno a TEI può condurre facilmente ad un utilizzo veloce ed efficiente di testi epigrafici codificati in TEI [22].

Gli approcci TEI e EpiDoc sono già stati adottati da molti progetti epigrafici [3], quali il progetto Dêmos (Università di Furman) e il corpus delle iscrizioni dei Macedoni e dei Traci sviluppato al KERA (centro di ricerca per l'antichità greca e romana di Atene) [22].

Anche altri studiosi considerano EpiDoc una scelta soddisfacente ed adeguata. Felle ha comparato [16] il progetto EAGLE (Electronic Archive of Greek and Latin Epigraphy¹¹) alle risorse sistematiche annotate con EpiDoc, arrivando alla conclusione che questi due tipi di risorse sono diversi ma complementari.

Alvarez et al. [53] e Gomez et al. [18] hanno discusso la possibilità di condividere l'informazione epigrafica come linked data basati su EpiDoc e hanno descritto l'implementazione della soluzione di linked data ideata per la loro base di dati *Hispania Epigraphica*.

Cayless [4] considera EpiDoc un rilevante strumento digitale per l'epigrafia, in grado di fornire una rappresentazione uniforme dei metadati epigrafici.

In conclusione, le linee guida TEI EpiDoc stanno emergendo come uno standard per l'epigrafia digitale. EpiDoc certamente non è l'unico modo possibile per utilizzare TEI per i testi epigrafici, ma gli strumenti, la ricca documentazione e i numerosi esempi¹² lo rendono un ottimo ambiente per progetti di nuove digitalizzazioni come quello qui descritto.

4.5 La struttura EpiDoc

Un documento EpiDoc è strutturato come un documento TEI standard con l'elemento `teiHeader` contenente alcune sezioni `Desc` iniziali (`fileDesc`, `encodingDesc`, `profileDesc`, `revisionDesc`, etc) per i metadati e per informazioni e descrizioni generali (in queste sezioni sono state annotate informa-

¹⁰La pagina web principale del progetto TEI è disponibile all'URL: <http://www.tei-c.org/index.xml>

¹¹Si consulti la pagina web del progetto all'URL: <http://www.eagle-eagle.it/>

¹²Vari esempi sono reperibili online: http://wiki.tei-c.org/index.php/Samples_of_TEI_texts

zioni quali: luogo di rinvenimento, periodo, tipo di supporto e ID degli oggetti e dei frammenti specifici). Un uso interessante della sezione `encodingDesc` è mostrato nella Figura 1: l'elemento `glyph` deve essere definito dentro il suo elemento genitore `charDecl`, a sua volta definito dentro l'elemento genitore del genitore `encodingDesc`.

Questa sezione, contenente una serie di elementi `charDecl`, è specializzata nella gestione delle informazioni specifiche relative alla codifica di glifi o caratteri particolari.

L'elemento `teiHeader` è poi seguito dall'elemento `text` contenente il `body`, a sua volta composto da una serie di `<div>` non numerati, distinguibili dai loro attributi `type` (un esempio dell'elemento `EpiDoc` `<div>` è mostrato nella Figura 2).

Divisioni tipiche possono includere:

- il testo stesso (`type="edition"`);
- una traduzione (`type="translation"`);
- una descrizione (`type="description"`);
- un commento (`type="commentary"`);
- informazioni storiche (`type="history"`);
- la bibliografia (`type="bibliography"`).

La DTD `EpiDoc` introduce un set finito di possibili valori per il tipo di `<div>`, in modo tale che ci sia una chiara distinzione tra sezioni riguardanti aspetti diversi, quali il commento la descrizione o la storia archeologica.

Un vantaggio dell'annotazione strutturata è che gli annotatori possono codificare molto meglio il proprio grado di certezza relativa all'annotazione di una specifica caratteristica. La data di un'iscrizione, ad esempio, può essere codificata come un intervallo di date possibili. `EpiDoc` include l'elemento TEI `<certainty>` e l'attributo `cert` per incoraggiare gli annotatori a specificare se essi siano o meno completamente sicuri di una specifica lettura o interpretazione data. Dopo alcune discussioni, la comunità `EpiDoc` ha deciso [22] che la "certezza" dovesse essere espressa come un valore Si/No, se l'annotatore è certo di una lettura oppure no. La certezza graduale è troppo complicata da gestire ed è probabilmente meglio spiegabile nel commento.

5 Sviluppo del corpus del Lineare A

La speranza che approcci computazionali possano aiutare a decifrare il Lineare A, insieme con l'evidente assenza di ricche risorse digitali in questo campo, hanno portato alla decisione di sviluppare questa nuova risorsa. In questa sezione sono descritte le problematiche riscontrate e le soluzioni adoperate.

```

<div lang="minoan"
  n="text"
  type="edition"
  part="N"
  sample="complete"
  org="uniform">
<head lang="eng">Edition</head>
<cb rend="front" n="HM 1673"/>
<ab part="N">
  <lb n="1"/>
  <w part="N">𐀀𐀁𐀂</w>
  <space dim="horizontal"
    extent="1em"
    unit="character"/>
  <w part="N">𐀃</w>
  <lb n="2"/>
  <w part="N">𐀄</w>
  <g ref="#n5"/>
  <w part="N">𐀅𐀆𐀇</w>
  <lb n="3"/>
  <w part="N">𐀈</w>
  <g ref="#n12"/>
  <w part="N">𐀉𐀊𐀋</w>
  <lb n="4"/>
  <w part="N">𐀌𐀍𐀎</w>
  <g ref="#n6"/>
  <lb n="5"/>
  <w part="N">𐀏𐀐𐀑</w>
  <lb n="6"/>
  <g ref="#n4"/>
  <w part="N">𐀒</w>
  <supplied reason="damage">
    ↓
  </supplied>
  <gap extent="2em"
    reason="lost"
    unit="character"
    dim="right"/>
</ab>
</div>

```

Figura 2: Esempio di un elemento EpiDoc <div> con attributo type="edition".

5.1 Raccolta dei dati

Fortunatamente l'esistenza del sito web di John G. Younger e dei volumi GORILA, insieme con l'Indice Raison-Pope, ha reso possibile un processo di raccolta semi-automatica, partendo dalle trascrizioni fonetiche/sillabiche prese dal sito web di John G. Younger (col suo esplicito permesso concesso via email), convertendole poi, per mezzo di script Python, in stringhe Unicode e importando nel contempo i metadati forniti in insieme con le trascrizioni (ID del luogo e del tipo di supporto, luogo di conservazione, periodizzazione etc.).

Le risorse di John G. Younger sul suo sito web consistono principalmente di due pagine HTML, una contenente iscrizioni originarie di Haghia Triada (che è il sito archeologico più ricco in termini di documenti rinvenuti) [50] e l'altra contenente documenti originari di tutte le altre località [51].

Le trascrizioni sono ben arricchite di metadati. I metadati forniscono la stessa informazione riportata nei volumi GORILA, incluso l'Indice Raison-Pope, più qualche descrizione aggiuntiva del supporto (descrizioni che sui volumi GORILA non erano necessarie, essendo le trascrizioni adiacenti alle immagini dei documenti) e il riferimento allo specifico volume e alla specifica pagina di GORILA.

5.2 Problemi di Segmentazione

Lavorando su antichi sistemi di scrittura, ci si imbatte spesso nel problema della segmentazione. John G. Younger spiega [42] che in Lineare A la separazione tra "parole", o meglio tra *gruppi di segni*, è indicata principalmente in due modi: in primo luogo, considerando i gruppi di segni con numeri o logogrammi come entità indipendenti con valore non morfemico, e dunque assumendo implicitamente che ci sia una separazione prima e dopo questi tipi di segni; in secondo luogo, posizionando un puntino tra due gruppi di segni o tra un gruppo di segni e un segno singolo di tipo numerico o logografico (ossia i due tipi di segni considerati "parole a sé" del primo punto), dunque separando esplicitamente i gruppi di segni tra loro.

John G. Younger spiega anche che, a maggior ragione in testi contententi lunghe catene di gruppi di segni, i puntini sono sempre usati per separare i gruppi di segni; questa caratteristica è particolarmente evidente in testi non burocratici e di tipo religioso.

Sul suo sito web, John G. Younger affronta anche il problema della sillabazione (che si manifesta come riporto a capo in caso di fine riga) [43], spiegando che in alcuni casi si nota una divisione su righe diverse di alcuni gruppi di segni. La ragione di ciò potrebbe avere a che fare con la separazione tra affissi di qualche tipo e le radici dei gruppi di segni. Come Younger stesso mette in chiaro, questa ipotesi richiederebbe una prova del fatto che siano coinvolti affissi nella morfologia del minoico. Il problema della sillabazione

è più complicato da risolvere perché una risorsa "neutrale" dovrebbe evitare trascrizioni che implicino una segmentazione ben conosciuta e definita per i gruppi di segni del Lineare A. Nelle trascrizioni di John G. Younger i gruppi di segni divisi su più righe sono riuniti, al fine di rendere chiaro quando uno specifico gruppo di segni, magari già noto, sia presente. Al contrario, il corpus digitale nel formato sopra descritto consente di mantenere fedelmente il testo, mentre tutta l'informazione relativa ad interpretazioni di questo tipo può essere conservata separatamente.

5.3 Trascrizioni Unicode

Sono state ottenute trascrizioni codificate in Unicode convertendo automaticamente le trascrizioni fonetiche/sillabiche di John G. Younger in trascrizioni con codifica GORILA (manualmente confrontate con i volumi GORILA) e poi convertendo automaticamente le trascrizioni con codifica GORILA in codici Unicode e stampandoli come caratteri Unicode (codifica UTF-8). Per creare i dizionari sillabe-GORILA e GORILA-Unicode, sono stati presi in considerazione la tabella di conversione di John G. Younger menzionata nella Sottosezione 3.4 e la documentazione ufficiale di Unicode (contenente esplicite informazioni di mappatura Unicode-GORILA). Tutte queste fasi di elaborazione sono state implementate con script Python.

5.4 Annotazione XML

Una volta raccolto l'intero corpus codificato in Unicode, è stata aggiunta automaticamente parte dell'annotazione XML per mezzo di script Python. Questi documenti devono poi essere manualmente corretti e completati confrontandoli con i volumi GORILA.

5.5 Un nuovo font del Lineare A

Prima del rilascio di Unicode 7.0, non c'era alcun modo di visualizzare i caratteri Unicode nell'intervallo 10600–1077F. Persino adesso, sistemi non aggiornati potrebbero avere problemi a visualizzare quei caratteri. Alcune implementazioni per il supporto Unicode in alcuni contesti (ad esempio per l'output \LaTeX) non sono sempre aggiornate, quindi non è affatto ovvio che siano disponibili font per i set di caratteri più recenti. Per questa ragione è stato scelto di sviluppare un nuovo font del Lineare A, risolvendo il problema principale riscontrato in `LA.ttf` (posizioni Unicode arbitrarie non corrispondenti con il set di caratteri rilasciato per il Lineare A). A partire dalla documentazione ufficiale di Unicode, è stato creato un set di simboli graficamente simili a quelli ufficiali, che sono poi stati allineati alle corrette posizioni Unicode. Il font è stato chiamato `John_Younger.ttf` come dimostrazione di stima e di gratitudine per il grande lavoro di John G. Younger, che ha reso accessibili al pubblico le informazioni di GORILA tramite mezzi

digitali. Con lo stesso scopo in mente è partita l'idea di sviluppare e distribuire liberamente questo nuovo font. Il font è disponibile alla pagina web <http://openfontlibrary.org/en/font/john-younger>.

6 Il corpus digitale del Lineare A verso gli open linked data

Come discusso da Chiarcos et al. [5] e da Oltramari et al. [26], modellare e pubblicare risorse linguistiche come linked data offre vantaggi significativi. Questa procedura può migliorare l'integrazione e l'interoperabilità delle risorse linguistiche. In questo formato, infatti, gli elementi linguistici sono univocamente identificabili, rendendo quindi possibile la navigazione dinamica di varie risorse interconnesse.

In questa prospettiva, il primo principio dei linked data proposto da Tim Berners Lee [2] considera l'URI (*Uniform Resource Identifier*) come il mezzo attraverso il quale le risorse possono essere univocamente identificate e che inoltre consente di ottenere informazioni su di esse. Questo ha portato alla scelta di fornire ID univoci per ciascun documento, tenendo in considerazione le convenzioni e gli standard *de facto* (Indice Raison-Pope) per riferirsi ai luoghi di rinvenimento, ai tipi di supporto e alla periodizzazione. Sono stati esplicitati anche i riferimenti alle fonti principali contenenti trascrizioni manuali, disegni/schizzi e foto delle iscrizioni, citando volumi e pagine specifiche di GORILA che è già il sistema di riferimento convenzionale usato dagli studiosi anche per le raccolte cartacee. D'altronde, gli studiosi del Lineare A, archeologi, filologi e linguisti storici, coinvolti nello studio del Lineare A applicavano già questo principio, facendo sempre riferimento agli indici e ai sistemi riferimento convenzionali.

Un secondo principio postulato da Tim Berners Lee è di basarsi sul protocollo HTTP, che è richiesto sia dai server web che dai browser e che può fornire specifici documenti, in diverse versioni o viste, in base a quale sia l'User Agent a richiedere l'URI. Per seguire questo principio, sono stati sviluppati documenti XML facilmente leggibili e trattabili anche automaticamente da software e script oltre che facilmente trasformabili in pagine HTML per mezzo di fogli di stile XSL. Fogli di stile diversi potrebbero fornire diverse strutture HTML, rispondendo di volta in volta alle varie necessità, in modo simile a come HTTP risponde con versioni e viste diverse del documento, identificato dalla URI, in base all'User Agent in questione.

Il terzo principio messo in evidenza da Tim Berners Lee richiede l'utilizzo di modelli di dati standardizzati e interoperabili per rappresentare e mettere a disposizione i linked data. I documenti XML assolvono questo requisito fornendo risorse interscambiabili e intercompatibili come spiegato nella Sottosezione 4.3. Coerentemente con questo principio, anche la codifica

dei caratteri è stata standardizzata, codificando tutto il testo con caratteri Unicode.

Come quarto principio Berners Lee propone di fornire collegamenti ad altre fonti; i riferimenti alle pagine web di John G. Younger e ai volumi GORILA potrebbero essere un modo per rispondere a questo requisito. Il formato utilizzato, d'altra parte, permette intrinsecamente di collegare i documenti ad altre risorse esistenti che seguano i principi dei linked open data, e a tutte quelle che saranno rese disponibili in futuro.

In conclusione, se una "nuvola dei Linked Open Data linguistici" [6] è possibile, il corpus digitale del Lineare A potrebbe un giorno farne parte.

7 Il Corpus digitale del Lineare A come risorsa culturale

Come definito dall'Unione Europea [12] e dall'UNESCO [36], il significato della nozione di *patrimonio culturale* non si applica solo a oggetti materiali e opere d'arte, ma anche al "patrimonio intangibile", come le tradizioni e le espressioni creative. In questa definizione rientrano di diritto anche i tipi di informazione conservati nei corpora linguistici; infatti, essi contengono informazioni riguardo alle tradizioni, alla conoscenza e allo stile di vita di una certa cultura.

Nonostante il minoico non sia stato ancora decifrato, il corpus del Lineare A fornisce interessanti informazioni riguardo all'economia, al commercio e alla religione.

Come accennato nella Sottosezione 3.1, Schoep ha fatto un'analisi critica delle tavolette e del loro ruolo nei processi amministrativi, basandosi sull'osservazione dei supporti fisici [35].

Ruth Palmer ha effettuato uno studio approfondito [30] della distribuzione dei beni (elencando precise quantità e luoghi) tra i vari centri minoici, anche senza una piena comprensione del contenuto dei documenti. Come Palmer mette in chiaro, "the ideograms for basic commodities, and the formats of the Linear A texts are similar enough to their Linear B counterparts to allow valid comparison of the types and amounts of commodities which appear in specific contexts", ossia, gli ideogrammi per alcuni beni di base e il formato di alcuni documenti commerciali sono talmente simili ai corrispettivi del Lineare B da rendere possibile un'analisi delle quantità di beni che apparivano in contesti specifici. Quindi, è possibile avere un'idea dell'attività economica e una misura dell'entità e della complessità delle transazioni. Dalle tavolette in Lineare A è possibile inferire informazioni riguardo la gestione delle risorse e riguardo il sistema di amministrazione dei centri minoici.

Van den Kerkhof e Rem [37] hanno analizzato le formule di libagione minoiche: iscrizioni religiose su coppe, stoviglie e tavolette che erano utilizzate per le offerte di olio e di altre bevande sacre all'alba. Le sacerdotesse che si

occupavano dei rituali di libagione utilizzavano ogni tipo di utensili e spesso iscrivevano le loro formule sacre su questi oggetti. Circa una trentina di questi testi, integri o solo parzialmente conservati, è sopravvissuta su tavolette di libagione, stoviglie e vasi, iscritti in vari tipi di scritture a mano. Sono disponibili trascrizioni di queste iscrizioni religiose ad opera di Consani [7] e di John G. Younger [52] sul suo sito web. Come notato da Duhoux [11] le formule di libagione minoiche hanno una struttura fissa con alcuni elementi variabili. Infatti, alcuni studi [9] riguardo la sintassi del minoico sono stati realizzati proprio sfruttando questa caratteristica e osservando l'ordine dei gruppi di segni riscontrato in queste formule regolari. Sono state avanzate ipotesi sul tipo di bevanda offerta sulla base di un importante indizio, la presenza di ideogrammi simili all'oliva potrebbe suggerire che i Minoici usassero l'olio per le loro libagioni [37] (e di certo l'oliva e il suo derivato, l'olio, sono prodotti da sempre molto comuni in area mediterranea). Dietro a queste numerose tracce del patrimonio culturale minoico già disponibili, se ne nasconde un'altra enorme ugualmente conservata: la lingua minoica con le sue storie, ancora non raccontate, rappresentative della vita di una grande civiltà di cui, eppure, ancora si sa così poco. Rimane la speranza che contributi come quello qui descritto possano essere utili per la comunità scientifica e che il minoico, nella sua rappresentazione digitale, possa finalmente essere decifrato per mezzo di approcci computazionali ancora intentati.

8 Progetti per il futuro

Per il futuro, oltre che al controllo manuale e al completamento della risorsa in via di sviluppo, si potrebbe pensare a fogli di stile XSL per creare pagine HTML per un'usabile fruizione sul web. A seguire, tutti i dati saranno disponibili e pubblicati online. Un passo in avanti, dal punto di vista delle risorse e degli strumenti utili a disposizione, potrebbe essere sviluppare un'interfaccia web per l'annotazione e per l'arricchimento dinamico delle informazioni del corpus.

Nota relativa agli indirizzi URL

Si precisa che tutti gli indirizzi URL menzionati nella relazione (bibliografia inclusa) sono stati riconsultati in data odierna: 22-07-2015.

Riferimenti bibliografici

- [1] Barriocanal, Elena García, Zeynel Cebeci, Aydın Öztürk e Mehmet C. Okur: *Metadata and Semantic Research: 5th International Conference, MTSR 2011, Izmir, Turkey, October 12-14, 2011. Proceedings*. Communications in Computer and Information Science. Springer,

- 2011, ISBN 9783642247309. https://books.google.it/books/about/Metadata_and_Semantic_Research.html?id=ydHP9iZRsd0C&hl=en.
- [2] Bizer, Christian, Michael T. Heath e Tim Berners Lee: *Linked data – the story so far*. Int. J. Semant. Web Inf. Syst. (IJSWIS), 2009, pagine 1–22.
- [3] Bodard, Gabriel: *EpiDoc: Epigraphic documents in XML for publication and interchange*. Latin on Stone: Epigraphic Research and Electronic Archives, Roman Studies: Interdisciplinary Approaches, forthcoming, 2009.
- [4] Cayless, Hugh: *Tools for Digital Epigraphy*. Proc. of the Association for Computing in the Humanities/Association for Literary and Linguistic Computing, Athens GA, 2003.
- [5] Chiarcos, Christian, John McCrae, Philipp Cimiano e Christiane Fellbaum: *Chapter 2 Towards Open Data for Linguistics: Linguistic Linked Data*. Theory and Applications of Natural Language Processing. Springer Berlin Heidelberg, 2013, ISBN 9783642317811, <https://books.google.it/books?id=4SZEAAAQBAJ>.
- [6] Chiarcos, Christian, Sebastian Nordhoff e Sebastian Hellmann: *Linked Data in Linguistics. Representing Language Data and Metadata*. 2012.
- [7] Consani, Carlo, Mario Negri, Francesco Aspesi e Cristina Lembo: *Testi Minoici Trascritti*. Rome: CNR–Istituto per gli studi micenei ed egeo-anatolici, 1999.
- [8] Cotterell, Arthur: *The Minoan World*. Charles Scribner’s Sons, 1980.
- [9] Davis, Brent: *Syntax in Linear A: The Word-Order of the ‘Libation Formula’*. Kadmos, 2014, ISSN (ONLINE) 1613-0723, (PRINT) 0022-7498.
- [10] Dietrich, Manfred e Oswald Loretz: *In memoriam Cyrus H. Gordon*. Ugarit-Forschungen. Ugarit-Verlag, 2001, ISBN 9783934628007. <https://books.google.it/books?id=tlrvMQEACAAJ>.
- [11] Duhoux, Yves: *Le Linéaire A: problèmes de Déchiffrement in Problems in Decipherment*. Bibliothèque des Cahiers de l’Institut de Linguistique de Louvain, 1989, pagine 59–119.
- [12] European Commission: *Supporting cultural heritage - European Commission*, 2015. http://ec.europa.eu/culture/policy/culture-policies/cultural-heritage_en.htm.
- [13] Evans, Arthur John: *Scripta minoa, vol. I*. Oxford: Clarendon Press, 1909.

- [14] Facchetti, Giulio M. e Mario Negri: *Creta minoica: sulle tracce delle più antiche scritture d'Europa*. LS Olschki, 2003.
- [15] Faure, Paul: *8. La Marle (Hubert), Linéaire A. La première écriture syllabique de Crète. Vol. 1: Essai de lecture; Vol. 2: Éléments de grammaire*. Revue des Études Grecques, 1998, pagine 339–340.
- [16] Felle, Antonio Enrico: *Esperienze diverse e complementari nel trattamento digitale delle fonti epigrafiche: il caso di EAGLE ed EpiDoc*. Diritto romano e scienze antichistiche nell'era digitale. Convegno di studio (Firenze, 12-13 settembre 2011), 2011, pagine 47–54.
- [17] Godart, Louis e Jean Pierre Olivier: *Recueil des inscriptions en linéaire A*. Librairie Orientaliste Paul Geuthner, 1976.
- [18] Gómez Pantoja, Joaquín L. e Fernando Luis Alvarez: *From relational databases to Linked Data in Epigraphy: Hispania Epigraphica Online, Elena García Barriocanal, Zeynel Cebeci, Mehmet C. Okur and Aydin Öztürk (eds.), Metadata and Semantic Research (Proceedings of the 5th International Conference, MTSR 2011, Izmir, Turkey, October 12-14, 2011)*. Communications in Computer and Information Science, 2011.
- [19] Hallager, Erik: *The Minoan roundel and other sealed documents in the neopalatial Linear A administration*, volume 14. Université de Liège, Histoire de l'art et archéologie de la Grèce antique, 1996.
- [20] Hunt, A.S.: *A note on the transliteration of papyri*. Nel *Chronique d'Égypte* 7, page 272–274, 1932, page 272–274.
- [21] Janko, Richard: *A stone object inscribed in Linear A from Ayios Stephanos, Laconia*. Kadmos, 1982, page 97–100.
- [22] Mahoney, Anne: *Electronic Textual Editing: Epigraphy*. TEI-Text Encoding Initiatives., 2007. http://www.tei-c.org/About/Archive_new/ETE/Preview/mahoney.xml.
- [23] Müller, Walter: *Die Tonplomben und andere gestempelte Tonobjekte*. I. Pini (ed.), Iraklion Archaeologisches Museum, Teil 6, Die Siegelabrücke von Aj. Triada und anderen Zentral- und Ostkretischen Fundorten. CMS (Corpus der Minoischen und Mykenischen Siegel) II,6. Berlin, 1999, pagine 339–400.
- [24] Müller, Walter: *Untersuchungen zur Typologie, Funktion und Verbreitung der Tonplomben von Knossos*. I. Pini (ed.), Iraklion Archäologisches Museum, Teil 8, Die Siegelabrücke von Knossos. CMS (Corpus der Minoischen und Mykenischen Siegel) II,8. Berlin, 2002, page 24–93.

- [25] Nagy, Gregory: *Greek-like Elements in Linear A*. Greek, Roman, and Byzantine Studies, 1963, pagine 181–211.
- [26] Oltramari, Alessandro, Piek Vossen, Lu Qin e E. Hovy: *New Trends of Research in Ontologies and Lexical Resources: Ideas, Projects, Systems*. Theory and Applications of Natural Language Processing. Springer Berlin Heidelberg, 2013, ISBN 9783642317811, <https://books.google.it/books?id=4SZEAAAAQBAJ>.
- [27] Owens, Gareth: *The structure of the Minoan language*. Journal of Indo-European Studies, 1999, pagine 15–55.
- [28] Packard, David W.: *Minoan Linear A*. University of California Press, 1974, ISBN 9780520025806. <https://books.google.it/books?id=vax3kwoscWQC>.
- [29] Palmer, Leonard Robert: *Luvian and Linear A*. Transactions of the Philological Society, 1958, pagine 75–100.
- [30] Palmer, Ruth: *Linear A commodities: a comparison of resources*. POLITEIA. Society and State in the Aegean Bronze Age, Aegaeum, 1995, pagine 133–156.
- [31] Raison, Jacques e Maurice Pope: *Index du linéaire A.*, volume 41 della serie *Incunabula Graeca*. Edizioni dell'Ateneo, 1971. <https://books.google.it/books?id=N10eYAAACAAJ>.
- [32] Raison, Jacques e Maurice Pope: *Corpus transnuméré du linéaire A*. BCILL (Louvain-la-Neuve). Isd, 1994, ISBN 9789068315615. <https://books.google.it/books?id=TbgcAQAIAAJ>.
- [33] Renfrew, Colin: *A Linear A tablet fragment from Phylakopi in Melos. (with a note on the inscription by William C. Brice)*. Kadmos, 1977, pagine 111–119.
- [34] Robinson, Andrew: *Writing and script: a very short introduction*. Oxford University Press, 2009.
- [35] Schoep, Ilse: *The administration of neopalatial Crete: a critical assessment of the Linear A tablets and their role in the administrative process*. Minos: Revista de filología egea, 2002, pagine 1–230.
- [36] UNESCO: *What is meant by "cultural heritage"?*, 2003. <http://www.unesco.org/culture/ich/index.php?lg=EN&pg=00022>.
- [37] Van den Kerkhof, Astrid e Peter Rem: *The Minoan libation formulas*. 2007.

- [38] Van Groningen, B.A.: *De signis criticis in edendo adhibendis*, 1932, pagine 362–365.
- [39] Van Groningen, B.A.: *Projet d'unification des systèmes de signes critiques*, 1932, pagine 262–269.
- [40] Younger, John G.: *Linear A Texts in phonetic Transcription: 1. List of Linked Files, Concordance: Raison-Pope-GORILA signs (a Microsoft Word document)*, 2000. http://www.people.ku.edu/~jyounger/LinearA/RAISON-GORILA_SIGNS.doc.
- [41] Younger, John G.: *Linear A Texts in Phonetic Transcription: 10c. Place Names*, 2000. <http://www.people.ku.edu/~jyounger/LinearA/#10c>.
- [42] Younger, John G.: *Linear A Texts in Phonetic Transcription: 16. Word Separation*, 2000. <http://people.ku.edu/~jyounger/LinearA/#16>.
- [43] Younger, John G.: *Linear A Texts in Phonetic Transcription: 17. Hyphenization*, 2000. <http://people.ku.edu/~jyounger/LinearA/#17>.
- [44] Younger, John G.: *Linear A Texts in phonetic Transcription: 4. Conventions (bibliographical, epigraphical)*, 2000. <http://people.ku.edu/~jyounger/LinearA/#4>.
- [45] Younger, John G.: *Linear A Texts in Phonetic Transcription: 5. Basic Statistics*, 2000. <http://www.people.ku.edu/~jyounger/LinearA/#5>.
- [46] Younger, John G.: *Linear A Texts in Phonetic Transcription: 7c. The Documents*, 2000. <http://people.ku.edu/~jyounger/LinearA/#7>.
- [47] Younger, John G.: *Linear A Texts in Phonetic Transcription: A map of Crete showing the location of the cretan sites*, 2000. http://people.ku.edu/~jyounger/pix/LinearA_Crete.jpg.
- [48] Younger, John G.: *Linear A Texts in Phonetic Transcription: A map of Greece showing the location of the greek sites outside Crete*, 2000. http://people.ku.edu/~jyounger/pix/LinearA_Greece.jpg.
- [49] Younger, John G.: *Linear A Texts in phonetic Transcription: Homepage*, 2000. <http://www.people.ku.edu/~jyounger/LinearA/>.
- [50] Younger, John G.: *Linear A Texts in phonetic Transcription: HT (Haghia Triada)*, 2000. <http://people.ku.edu/~jyounger/LinearA/HTtexts.html>.
- [51] Younger, John G.: *Linear A Texts in phonetic Transcription: Other Texts (not Haghia Triada)*, 2000. <http://people.ku.edu/~jyounger/LinearA/misctexts.html>.

- [52] Younger, John G.: *Linear A Texts in Phonetic Transcription: Presumably Religious Inscriptions*, 2000. <http://www.people.ku.edu/~jyounger/LinearA/religioustexts.html>.
- [53] Álvarez, Fernando Luis, Elena García Barriocanal e Joaquín L. Gómez Pantoja: *Sharing Epigraphic Information as Linked Data*. Nel *Metadata and Semantic Research*. Springer, 2010, pagine 222–234.