

SCUOLA NORMALE SUPERIORE DI PISA
Laboratorio di Storia, Archeologia e Topografia del Mondo Antico

QUARTE
GIORNATE INTERNAZIONALI DI
STUDI SULL'AREA ELIMA

(Erice, 1-4 dicembre 2000)

ATTI

III

Pisa 2003

Il presente volume è stato curato da Alessandro Corretti.

ISBN 88-7642-122-X

SNS-GREEK & LATIN: SVILUPPI FUTURI

ANTONELLA RUSSO

SNS-Greek & Latin è un'applicazione per Macintosh sviluppata presso il Laboratorio Informatico per le Lingue Antiche della Scuola Normale Superiore. Il Laboratorio fu fondato nel 1991 da Giuseppe Nenci, che grazie alla sua lungimiranza riuscì fin dal 1987 a prevedere l'importanza dell'uso degli strumenti informatici nelle discipline classiche dando avvio alla realizzazione del progetto *SNS-Greek* e seguendolo con impegno in tutti i suoi sviluppi successivi. La versione attuale di *SNS-Greek & Latin* è la 5.1, la cui distribuzione è iniziata a fine novembre 2000. Essa consente di utilizzare tre banche dati su CD-ROM: il *Thesaurus Linguae Graecae* versione 'E', il PHI #5.3 e il PHI #7. Queste banche dati sono accomunate dal fatto di essere distribuite prive di software di consultazione e proprio per questa ragione la Scuola decise di realizzarne uno inizialmente per uso interno che poi, su richiesta di molti studiosi, venne diffuso in numerose istituzioni italiane ed estere. Attualmente *SNS-Greek & Latin* ha alcune centinaia di utenti distribuiti in più di 20 venti paesi nel mondo.

Il *Thesaurus Linguae Graecae* o TLG è una banca dati contenente testi letterari greci da Omero al 600 d. C. oltre a buona parte della documentazione posteriore fino alla caduta di Costantinopoli. Essa è prodotta e distribuita dall'Università di Irvine in California. Il progetto *Thesaurus Linguae Graecae* è iniziato nel 1972 sotto la direzione di Theodore Brunner ed è un progetto tuttora attivo in quanto l'equipe del TLG attualmente diretta da Maria Pantelia sta continuando ad immettere testi e sarà dunque presentata in futuro almeno una versione migliorata. La versione 'E' è uscita nel marzo 2000 e contiene 1837 autori, più

di 7000 opere per un totale di circa 600 milioni di caratteri di testo. Essa è codificata in Betaform, un particolare formalismo che fa corrispondere ad ogni segno alfabetico o diacritico greco un solo carattere *ascii* e che consente inoltre di codificare anche tutta la formattazione del testo in modo che sia possibile riprodurre fedelmente su computer l'edizione a stampa. Il TLG è la più estesa e la più completa delle banche dati che *SNS-Greek & Latin* consente di gestire: comprende infatti, oltre ai testi, un *Canone* elettronico di classificazione degli autori e delle opere e degli indici.

La seconda banca dati che l'applicazione consente di consultare è il PHI #5.3, prodotta dal Packard Humanities Institute di Los Altos e rilasciata nel 1992. Di questa banca dati non sono previsti ulteriori aggiornamenti. Essa contiene testi latini classici non cristiani dalle origini al II d. C. oltre a qualche testo posteriore ed è una banca dati multilingue in quanto contiene anche la Bibbia copta, la Bibbia ebraica e il Nuovo Testamento greco. Questa banca dati è decisamente più piccola del TLG, contiene infatti circa 350 autori con 95 milioni di caratteri di testo. Anch'essa è codificata in Betaform ma è meno ricca di informazioni in quanto non contiene né un *Canone* di classificazione, né indici, contiene solo un file con indicato per ogni testo, l'edizione usata.

L'ultimo CD-ROM utilizzabile da *SNS-Greek & Latin* è il PHI #7 che, prodotto dal Packard Humanities Institute, contiene due banche dati: la Duke data bank of documentary papyri e iscrizioni greche, anche di epoca cristiana. Il PHI #7 è stato realizzato a cura di numerose università americane fra cui la Duke University e l'University of Michigan per i papiri e la Cornell e Ohio University per le iscrizioni. Questa banca dati è uscita nel 1997 e di sicuro ne uscirà una versione aggiornata in quanto le diverse università interessate continuano a immettere e a curare le versioni elettroniche delle raccolte. Al momento contiene 103 milioni di caratteri ed è dunque di dimensioni paragonabili all'altra banca dati del Packard; è codificata in Betaform e contiene due livelli di testo compresenti: testo greco e notazione infralinea, come da edizione a stampa, contenente vari tipi di informazione quali il sito di rinvenimento, la datazione, il tipo di

supporto, il tipo di scrittura e molta informazione di carattere bibliografico.

L'ultima versione di *SNS-Greek & Latin* è uscita a fine novembre 2000 e contiene due novità: l'utilizzo dei nuovi indici SNS-E del TLG e la possibilità di scegliere anche i font del gruppo Greekkeys¹ per visualizzare ed esportare i testi greci.

Numerose migliorie sono state apportate fra le quali la creazione di una nuova finestra di *append* per il TLG, l'attivazione del *goto* per papiri e iscrizioni, la gestione di testi copti contenuti nel PHI #5.3 e nel PHI #7 e alcune opzioni di esportazione.

La selezione dei testi nella banca dati TLG è ora possibile tramite due diverse maschere di scelta. La prima è quella tradizionale del *Canone*, mostrata in figura 1. È stata aggiunta una nuova voce di *append* che permette di selezionare per numero (*Append from...to...*) una gamma di autori. La seconda (in figura 2) permette di scegliere fra autori ordinati alfabeticamente ed è utile dato che non si devono ricordare i numeri associati agli autori. La

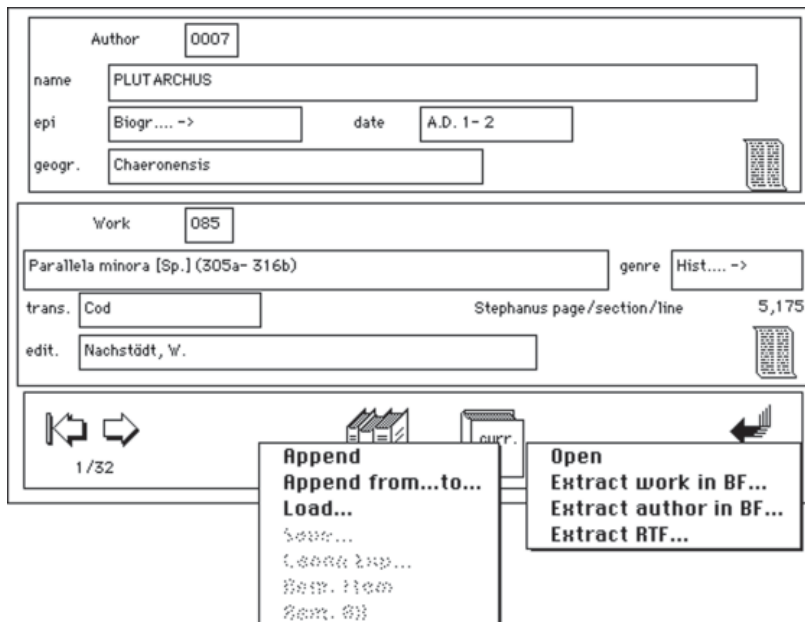


Fig. 1.

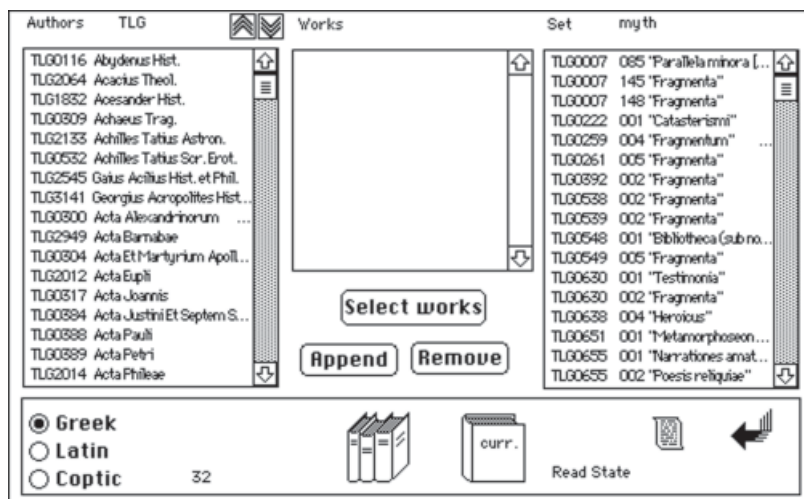


Fig. 2.

finestra è analoga a quella utilizzata per i due CD-ROM del Packard Humanities Institute. È possibile creare *set* nella finestra di *Canone*, tramite i criteri disponibili, quali genere letterario, secolo, e così via, salvarli con un nome e poi ricaricarli dalla finestra *TLG list* se si desidera aggiungere ulteriori autori senza ricorrere al numero.

La grande novità relativa alla versione 5.1 è la possibilità di utilizzare i nuovi indici SNS-E, prodotti e distribuiti su CD-ROM dal Laboratorio Informatico per le Lingue Antiche (LILA), allo scopo di rendere estremamente più veloce la ricerca estesa all'intera banca dati TLG.

A questo punto ci sono tre possibili metodi di accesso ai dati: sequenziale, il metodo più lento in cui tutti i testi devono essere interamente scanditi, tramite gli indici del TLG già presenti sul CD-ROM stesso e i nuovi indici SNS-E realizzati dal LILA. La differenza principale fra i due tipi di indici è la seguente: nei vecchi indici su CD TLG vi sono riferimenti generici all'opera mentre nei nuovi indici sono indicate le posizioni precise in cui la forma cercata ricorre. Questa differenza si riflette chiaramente sui tempi di ricerca.

Ad esempio la ricerca di un termine poco attestato come

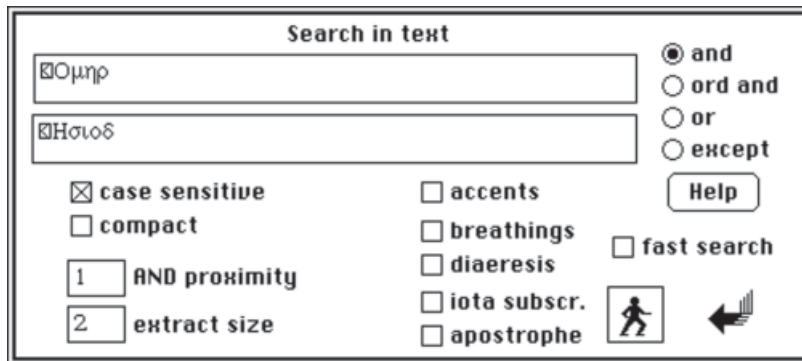


Fig. 3.

ΥΠΕΡΑΤΤΙΚ, che ricorre 13 volte in 11 opere diverse richiede, utilizzando i vecchi indici 6 minuti e 8 secondi su un PowerMac 8500. Devono essere infatti, rintracciate le opere dall'indice e poi ognuna di esse deve essere scandita sequenzialmente per trovare le occorrenze dato che non si ha sufficiente informazione sui punti ove esse ricorrono. Questa tecnica diventa tra l'altro inutilizzabile per termini molto attestati poiché i tempi di ricerca diventano paragonabili a quelli di una normale ricerca sequenziale su tutto il CD. Utilizzando invece i nuovi indici, che di certo richiedono più spazio ma sono ricchi di molta più informazione, la stessa ricerca impiega circa 1 secondo. Un po' di numeri relativi ai nuovi indici sono i seguenti: l'utente deve avere circa 837 milioni di bytes liberi sul proprio hard disk per poter scaricare gli indici, che contengono 1.087.953 forme distinte e 87 milioni di parole. Le linee di testo contenute nel TLG E ovvero le citazioni distinte sono 9.008.249.

In figura 3 è presentata la maschera di ricerca relativa ai nuovi indici. Essa è del tutto analoga alla maschera consueta di ricerca sequenziale e consente quasi tutte le opzioni presentate, ovvero la scelta degli operatori logici, l'attenzione alla maiuscola e agli accenti, spiriti, dieresi, iota sottoscritto e apostrofo oltre alla possibilità di utilizzare quasi tutti i metacaratteri (caratteri speciali indicatori di inizio e fine parola, di carattere indeterminato, di scelta di caratteri in una certa posizione, di esclusione di un certo carattere in una data posizione).

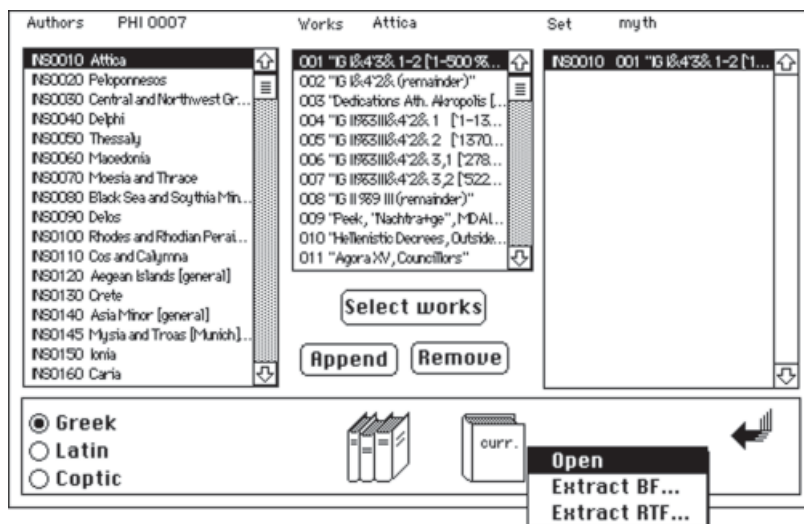


Fig. 4.

Un'altra nuova possibilità è quella di poter aprire un papiro documentario o iscrizione, contenuti nel CD-ROM PHI #7, e posizionarsi in un dato luogo. In figura 4 è mostrata la finestra da cui si può scegliere, ad esempio, una raccolta di iscrizioni e tramite la voce di menù *Open*, posizionarsi all'inizio della iscriz-

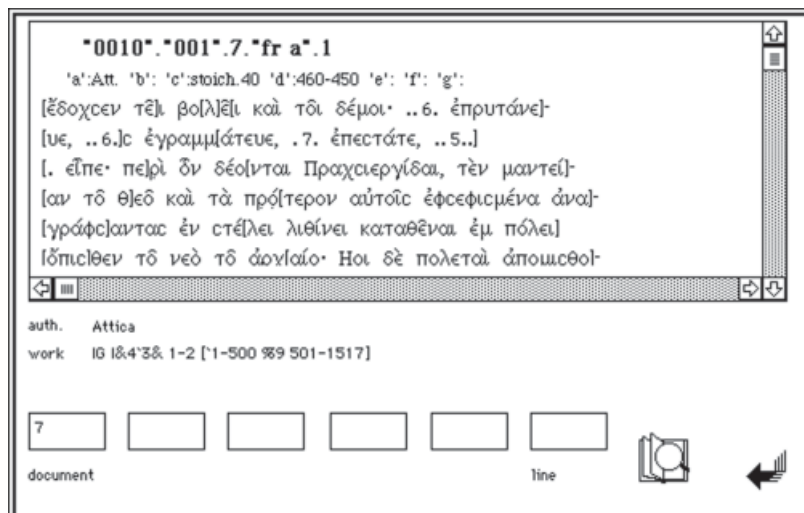


Fig. 5.

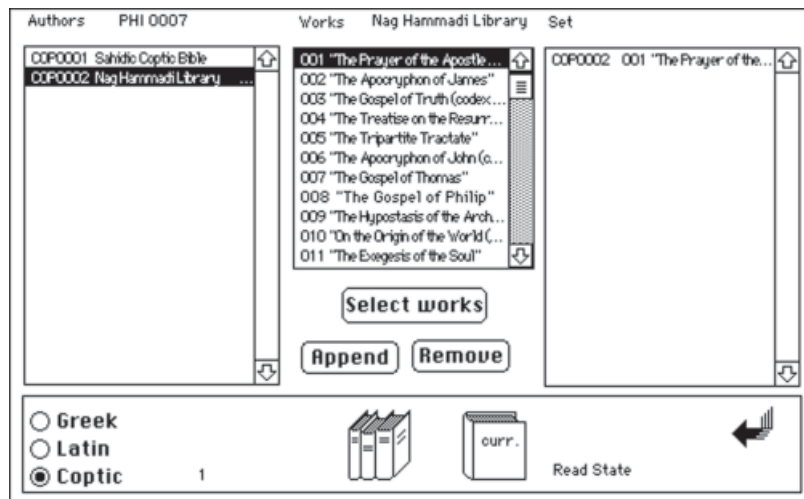


Fig. 6.

zione 7, come è indicato in figura 5. Questa possibilità esisteva già per i testi greci contenuti nel TLG e latini contenuti nel PHI#5.3 ed è stata ora estesa anche ai papiri documentari ed epigrafi contenuti nel PHI #7.

Le due banche dati prodotte dal Packard Humanities Institute contengono anche testi in altre lingue, in particolare testi copti. In passato era possibile solo accedere ai testi latini contenuti nel PHI #5.3 e ai testi greci contenuti nel PHI #7. Ora la maschera a liste di accesso ai testi è stata modificata in modo che l'utente possa scegliere una lingua (greco, latino o copto) e consultare i testi in quella lingua contenuti nel CD corrente. In particolare il PHI #5.3 contiene, oltre ai testi latini classici, il Nuovo Testamento greco, il testo dei Settanta e il Nuovo Testamento copto. Il PHI #7 contiene anche iscrizioni latine, la Bibbia copta e i testi della biblioteca di Nag Hammadi. Ad esempio in figura 6 è mostrata la maschera di scelta dei testi copti contenuti nel PHI #7. Dopo avere scelto un testo, ad esempio il primo della raccolta di Nag Hammadi, è possibile formulare una ricerca e quindi consultare i casi trovati (vedi fig.7) nel font copto CopticLS². Alcuni simboli del Betaform quali soprallineature o sottolineature e diversi altri, che non è stato possibile rendere nel font scelto, sono indicati in una finestra che compare facendo click sull'icona con il punto

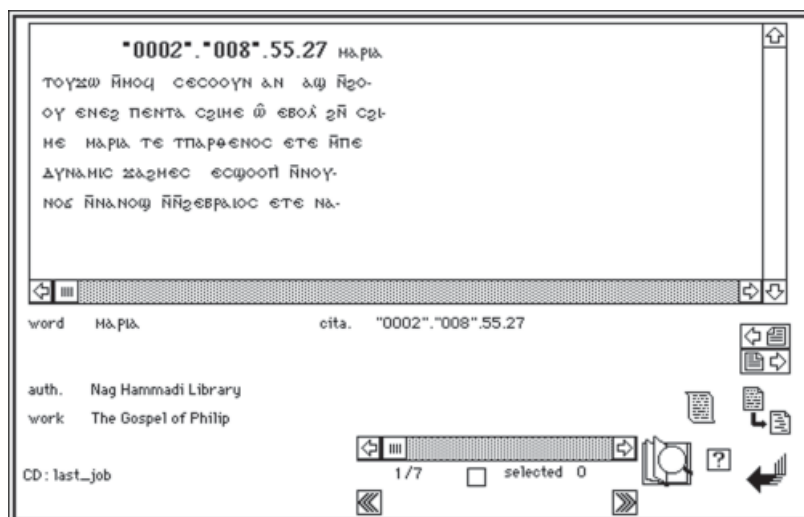


Fig. 7.

interrogativo, in un elenco dove accanto al simbolo del betaform (ad esempio %81) compare il suo significato. Tutti i casi trovati in testi copti possono essere esportati nel font CopticLS che, grazie ad un accordo del LILA con la Linguist's software, è messo a disposizione dell'utente nella sua versione bitmap 12 punti (così come il font GraecaII in versione bitmap 14 punti).

I progetti futuri del nostro Laboratorio vanno in tre direzioni: sviluppo di una versione per Windows di *SNS-Greek & Latin*, introduzione di nuove opzioni in *SNS-Greek & Latin* per Macintosh e produzione di software didattico.

La prima direzione è in stato avanzato di sviluppo poiché esiste già un prototipo che deve essere però ingegnerizzato prima di essere distribuito. La versione Windows avrà le stesse caratteristiche e possibilità offerte dalla versione Macintosh ed ha, per ora, una interfaccia molto simile (vedi figura 8 per un esempio di finestra di accesso alla banca dati TLG via Canone elettronico).

Nuove migliorie da apportare al programma *SNS-Greek & Latin* vengono continuamente richieste dagli utenti. In particolare la modifica dell'interfaccia in modo da adattarla alle maggiori dimensioni degli schermi dei monitor attuali, l'evidenziazione della riga o stringa cercata nei casi trovati e così via.

The screenshot displays a software interface for managing author and work data. It is organized into three main sections:

- Author Section:** Contains fields for 'Author' (0012), 'name' (HOMERUS), 'epi' (Epic), 'date' (8 B.C.), and 'geogr.'. A small icon is visible to the right of the 'geogr.' field.
- Work Section:** Contains fields for 'Work' (001), 'lines', 'genre' (Epic), 'trans.' (Cod), 'Book/line' (115,477), and 'edit.' (Allen, T.W.). A small icon is visible to the right of the 'edit.' field.
- Navigation Section:** Located at the bottom, it includes a '1/4' indicator, a 'Read State' label, and several icons for navigating through the data.

Fig. 8.

Per quanto riguarda il software didattico vi sono due prospettive: l'uso di *SNS-Greek & Latin* in questo campo e l'uso di *authoring software* (o programmi autore).

SNS-Greek & Latin consente attualmente di creare ed esportare indici di singoli testi (greci del TLG o latini del PHI #5.3), ordinati alfabeticamente. La possibilità di indicizzare frequenzialmente i testi (ma anche l'intera banca dati) consentirebbe la creazione di strumenti per un insegnamento intensivo della lingua greca(o latina). Inoltre un passo successivo che porterebbe a una qualità superiore nelle ricerche e nella indicizzazione potrebbe essere la lemmatizzazione degli indici.

L'uso di programmi autori, che consentono di preparare test molto flessibili utilizzando tipi di risposte diverse (multiple, hot spot, ecc), si rivela estremamente utile per fare esercizio e migliorare le conoscenze della lingua, della storia e della cultura greca e romana, per test di verifica o per creare percorsi di studio.

NOTE

¹ The American Philological Association, 291 Logan Hall, 249 S. 36th Street, The University of Pennsylvania, Philadelphia, PA 19104-6304 USA, E-mail: apaclassics@sas.upenn.edu.

² Linguist's Software Inc., PO Box 580, Edmonds, WA 98020-0580, USA; tel (425) 775-1130, fax (425) 771-5911; <http://www.linguistsoftware.com>.