THESARURUS LINGUAE SERICAE

Its History in Comparative Perspective, its Aspirations and its Implementation in FileMaker.

by Christoph Harbsmeier
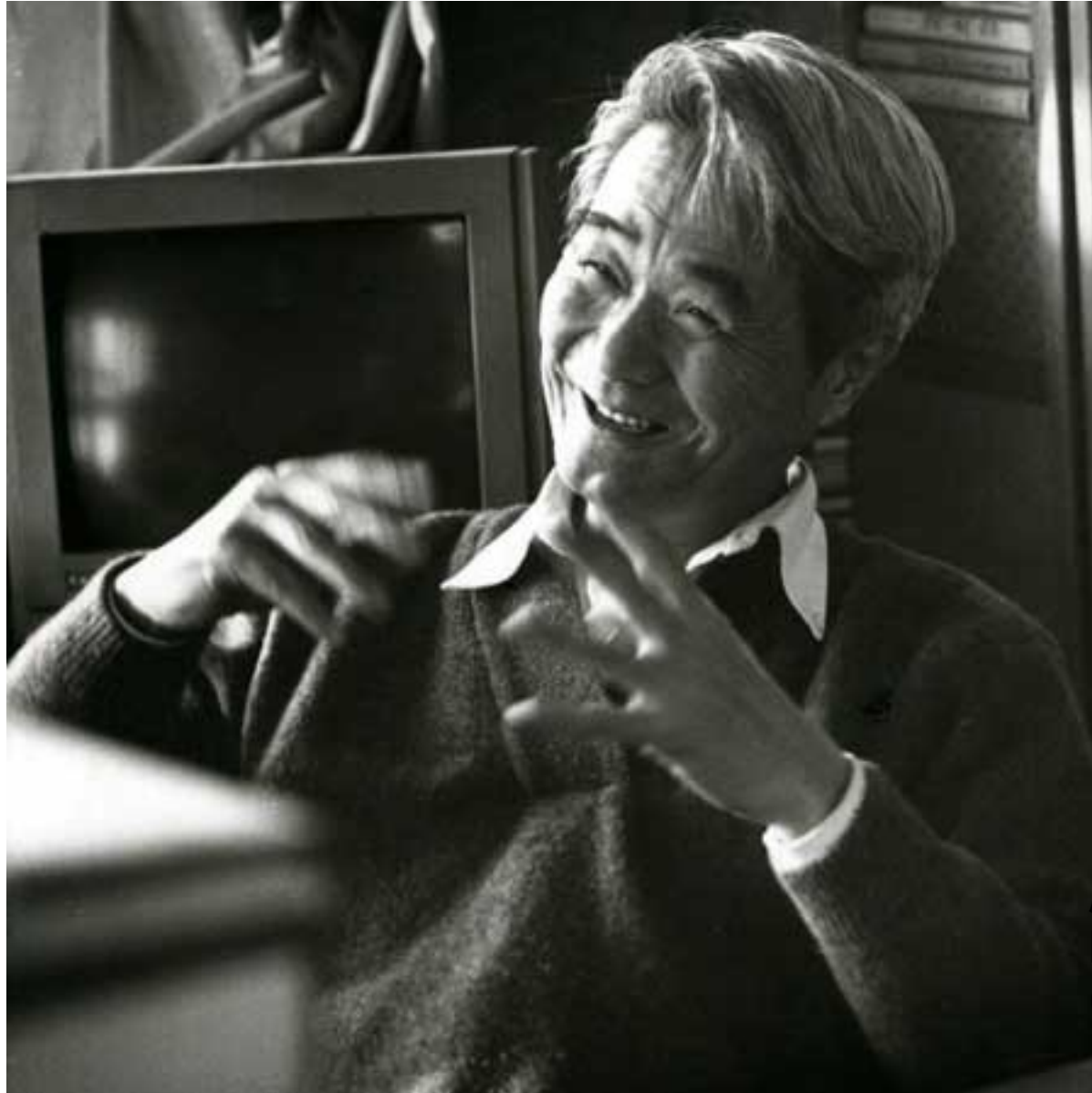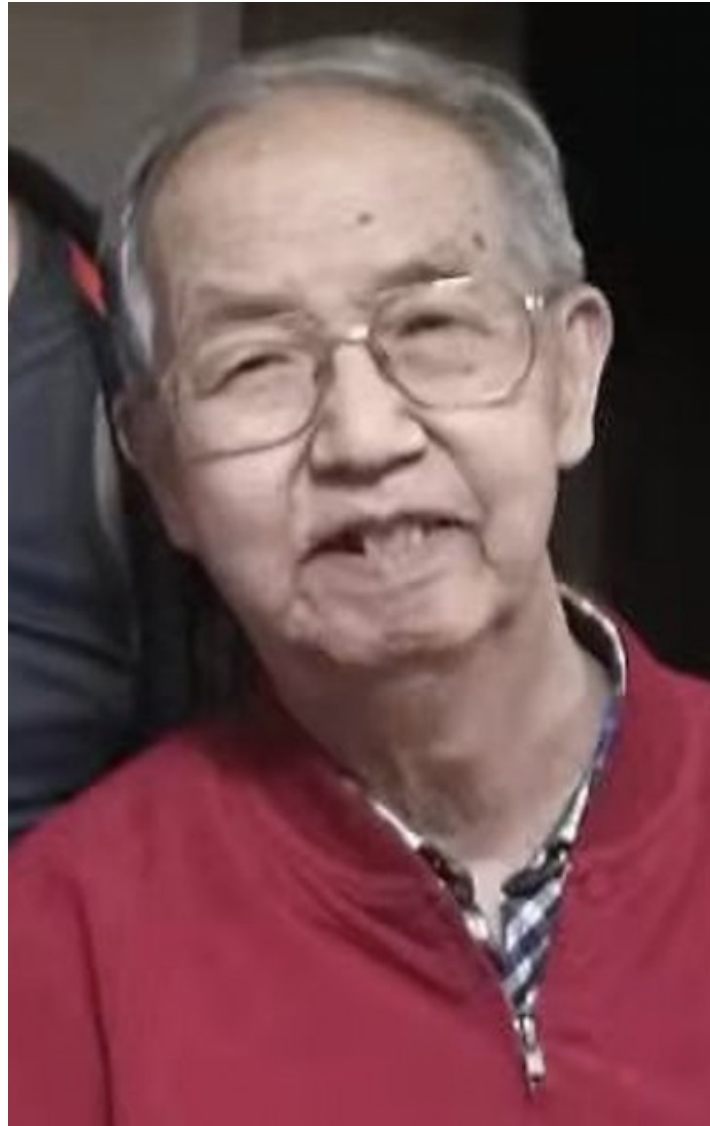
何莫邪

University of Oslo

港學文典

龐樸題

龐樸
Pang Pu
（1928-2015）

Jiang Shaoyu
蔣紹愚

Göran Malmqvist

NOVVS LINGVAE ET ERVDITIONIS ROMANAE THESAVRVS

POST
RO. STEPHANI
ET ALIORVM
NVPER ETIAM IN ANGLIA
ERVDITISSIMORVM HOMINVM
CVRAS
DIGESTVS, LOCVPLETATVS, EMENDATVS

ET

GEORGIO II

DEFENSORI FIDEI
ET
PACATORI ORBIS
BIBLIOTHECAM ACADEMIAE SVAE
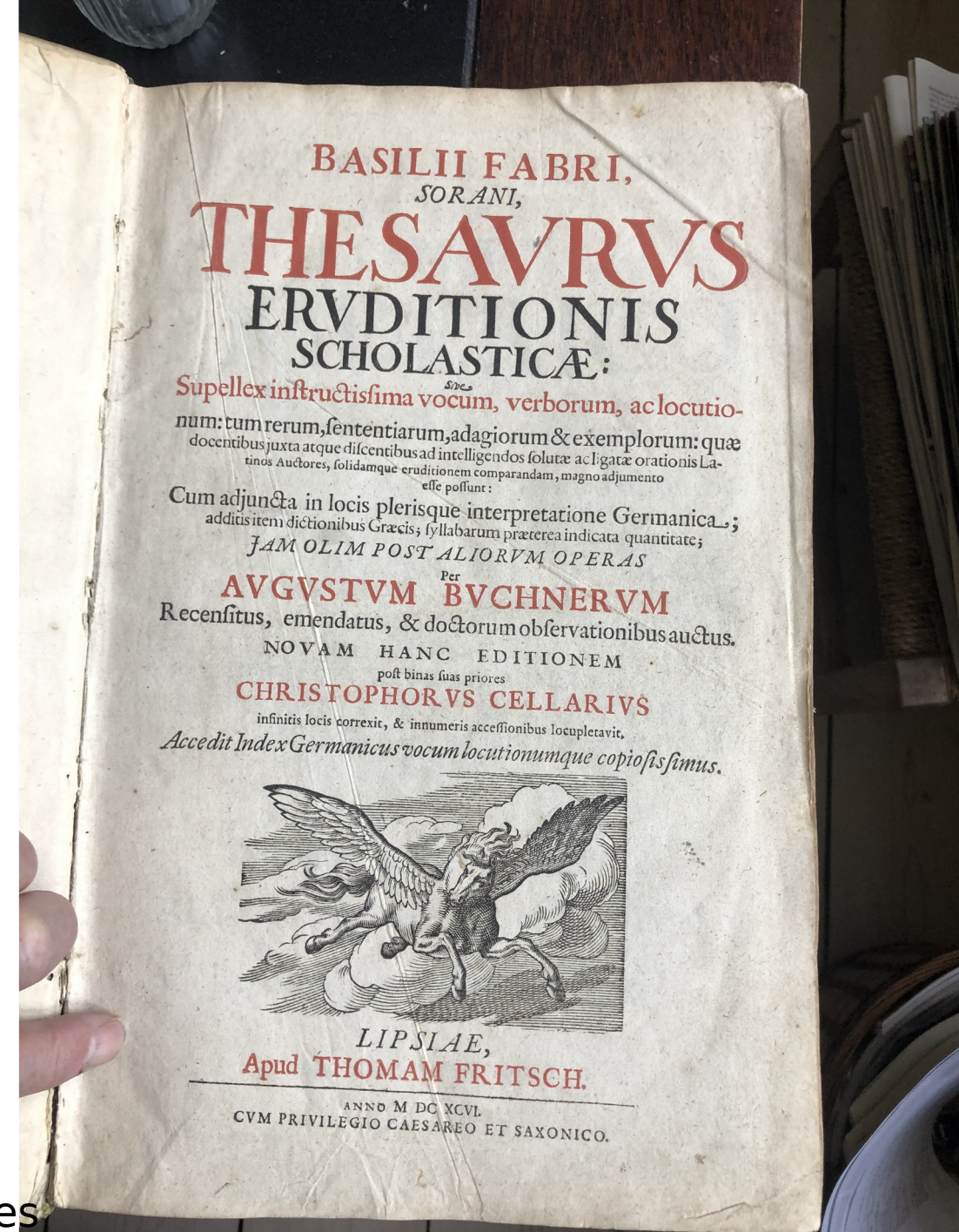GEORGIAE AVGVSTAE
QVAE GOTTINGAE EST
INVISENTI
DEVOTI ANIMI PIETATE OBLATVS
A. D. XXX. IVL. CIƆ IƆCC XXXXVIII.
A
IO. MATTHIA GESNERO
ELOQ. ET POES. P. P. O. ET ACAD. BIBLIOTHEC.

CVM PRIVIL. SACR. CAES. MAIEST. AC SER. REG. POL. ET EL. SAX.

LIPSIAE
IMPENSIS CASP. FRITSCHII VIDVAE ET BERNH. CHR. BREITKOPFII
M DCC XLIX.

Stephane Estienne
1503-1559

ed. 1641
2557 pages

BASILII FABRI, SORANI, THESAVRVS ERVDITIONIS SCHOLASTICÆ:

Sive Supellex instructissima vocum, verborum, ac locutionum: tum rerum, sententiarum, adagiorum & exemplorum: quæ docentibus juxta atque discentibus ad intelligendos solutæ ac ligatæ orationis Latinos Auctores, solidamque eruditionem comparandam, magno adjumento esse possunt:

Cum adjuncta in locis plerisque interpretatione Germanica; additis item dictionibus Græcis, syllabarum præterea indicata quantitate;

JAM OLIM POST ALIORVM OPERAS
Per
AVGVSTVM BVCHNERVM
Recensitus, emendatus, & doctorum observationibus auctus.
NOVAM HANC EDITIONEM
post binas suas priores
CHRISTOPHORVS CELLARIVS
infinitis locis correxit, & innumeris accessionibus locupletavit.
Accedit Index Germanicus vocum locutionumque copiosissimus.

LIPSIAE,
Apud THOMAM FRITSCH.
ANNO M DC XCVI.
CVM PRIVILEGIO CAESAREO ET SAXONICO.

BASILII FABRI,
SORANI,
THESAVRVS
ERVDITIONIS
SCHOLASTICÆ:

Sive

Supellex inſtructiſſima vocum, verborum, ac locutio-
num: tum rerum, ſententiarum, adagiorum & exemplorum: quæ
docentibus juxta atque diſcentibus ad intelligendos ſolutæ ac ligatæ orationis La-
tinos Auctores, ſolidamque eruditionem comparandam, magno adjumento
eſſe poſſunt:

Cum adjuncta in locis pleriſque interpretatione Germanica;
additis item dictionibus Græcis; ſyllabarum præterea indicata quantitate;

JAM OLIM POST ALIORVM OPERAS

Per

AVGVSTVM BVCHNERVM
Recenſitus, emendatus, & doctorum obſervationibus auctus.

NOVAM HANC EDITIONEM
poſt binas ſuas priores

CHRISTOPHORVS CELLARIVS
infinitis locis correxit, & innumeris acceſſionibus locupletavit.

Accedit Index Germanicus vocum locutionumque copioſiſſimus.

LIPSIAE,
Apud THOMAM FRITSCH.

ANNO M DC XCVI.

CVM PRIVILEGIO CAESAREO ET SAXONICO.

Thesaurus Linguae Latinae - Jaxtyn Leith

OXFORD LATIN DICTIONARY

EDITED BY
P. G. W. GLARE

SECOND EDITION

OXFORD
UNIVERSITY PRESS

THE CAMBRIDGE GREEK LEXICON

VOLUME I
A–I

What to learn from TLL and **Cambridge Greek Dictionary**: 24 year's work, 6 full-time collaborators

- 1. It is essential in TLS to specify the range of subjects, direct and indirect objects as well as adverbs, for verbs and typical adjectives and other modifiers for nouns. This needs a lot of work of revision, but new entries must follow the new requirement.

- 2. It is essential everywhere in TLS to begin definitions, wherever possible, with an analytic elucidation of the meaning before turning to suggestions for translation into English. TLS has been insufficiently systematic in this so far, but this is remediable.

- 3. Perhaps to indicate abbreviated indications of typical authors, where uses are found. This in fact, which gives a nice idea of history

- 4. Thesaurus Linguae Latinae has increasingly become more advanced in the analytic procedures used to relate systematically the separate meanings of words and derive them from root meanings.

- TLS has a HUGE AMOUNT to learn from TLL and CGD on the subtle and idiomatic description of meanings. In this area
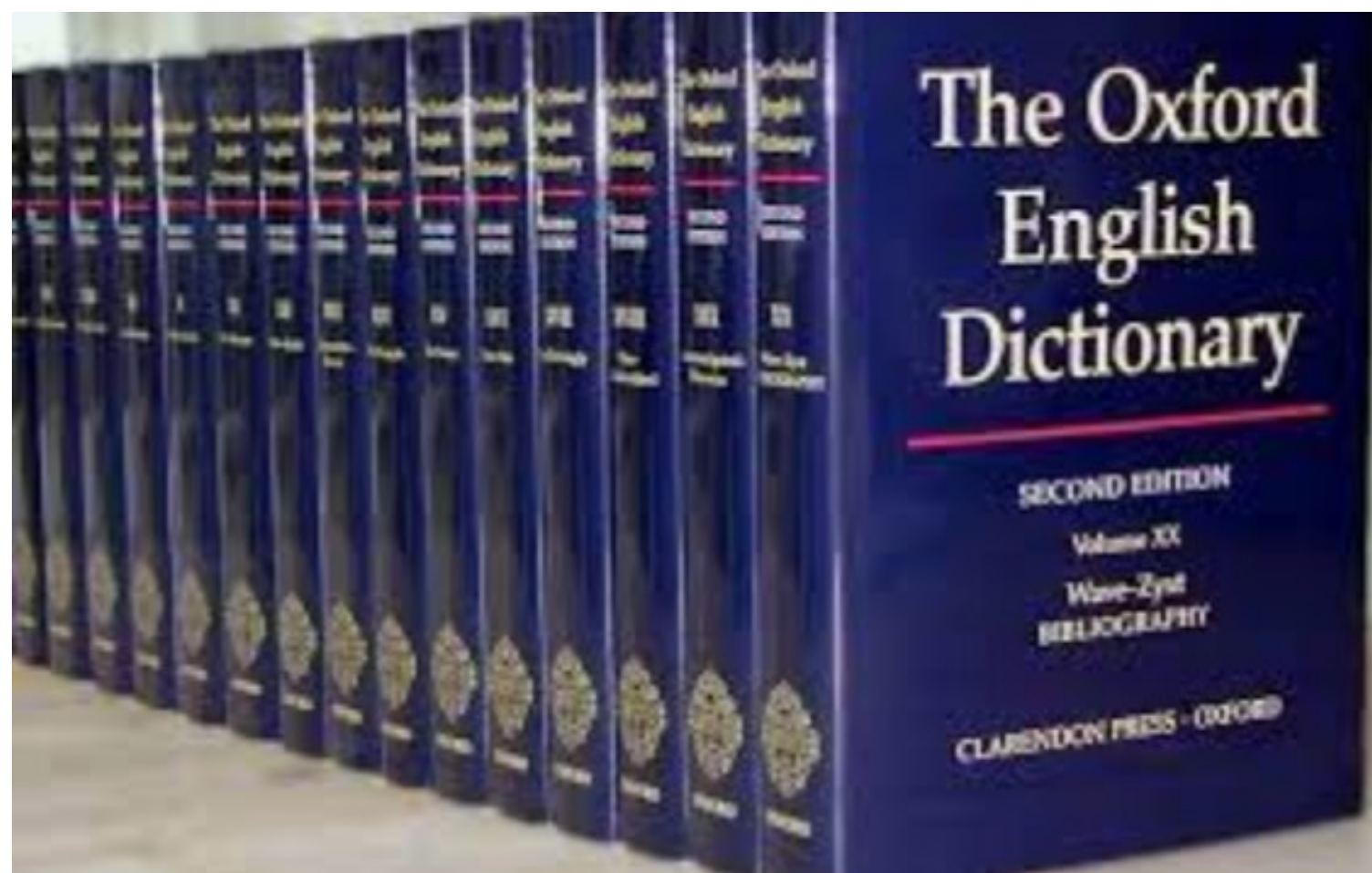
The Oxford English Dictionary and Thesaurus

- The OED is arguably the greatest lexicographic achievement in the world.
- By comparison, TLS fades into insignificance.
- TLS has achieved nothing.
- In what follows I shall discuss aspirations.

HISTORICAL THESAURUS OF THE OXFORD ENGLISH DICTIONARY

VOLUME I

HISTORICAL THESAURUS OF THE OXFORD ENGLISH DICTIONARY

VOLUME II

OXFORD

OXFORD

Historical Thesaurus

of the Oxford English Dictionary

with additional material from *A Thesaurus of Old English*

Christian Kay   Jane Roberts   Michael Samuels   Irené Wotherspoon

Volume I: **Thesaurus**

over 4 000 pages

OXFORD
UNIVERSITY PRESS

<div style="display:flex">

blue
n. 01.01.05.05
n. 01.01.10.03
n. 01.02.04.13.09.02.01 | 07.06
n. 01.02.04.13.09.02.05 | 01.04.02
n. 01.02.06.13.11.08.10 | 05.04.01
n. 01.02.06.20.01.01 | 12.06.04
n. 01.02.06.20.05.02.01 | 12
n. 01.02.08.01.09.01.02 | 08
n. 01.02.08.01.09.03.01 | 05.01
n. 01.02.08.04.12.05.10 | 07.01
n. 01.02.09.14.01.02 | 08
n. 01.02.10.02 | 20; 26.06.07
n. 01.02.11.01.01.05 | 09
n. 01.04.09.07.06
n. 01.04.09.07.06.03 | 02
n. 02.01.12.05 | 05; 05.01
n. 02.01.12.08.06.01.01 | 05
n. 03.01.07.03 | 01
n. 03.03.17.05.01 | 09.01
n. 03.03.17.06
n. 03.04.06.17.03 | 03.02.07
n. 03.04.13.10.01 | 09
n. 03.04.13.12.04 | 04
n. 03.05.05.07.03.01
n. 03.06.03.02 | 08.05
n. 03.06.03.02.02 | 19
n. 03.10.12.03.03.01.03 | 24.02
n. 03.11.01.05.02.07.04 | 07
n. 03.11.03.01.03 | 01
n. 03.11.04.13.05.02 | 09.01
adj. 01.02.01.01.07 | 04
adj. 01.02.05.09.05
adj. 01.02.06.12.01 | 19
adj. 01.02.08.02.22.01 | 04
adj. 01.02.09.05 | 10
adj. 01.02.10.03 | 17; 21
adj. 01.04.03.03.02.03.02 | 13.07
adj. 01.04.07.03.07.03 | 01.02
adj. 01.04.09.07.06 | 00; 05; 12; 13
adj. 01.04.09.07.09
adj. 02.01.12.05 | 02
adj. 02.01.14.06.01
adj. 02.01.15.07.09.52 | 08
adj. 02.02.30.08
adj. 02.05.05.02
adj. 03.04.06.17.03 | 03
adj. 03.04.06.17.05.01 | 01.01
adj. 03.05.05.07.03.01
adj. 03.11.03.01.06 | 09
vi. 02.02.29.03.01
vt. 01.02.11.01.01.02 | 10
vt. 01.04.09.07.06.02
vt. 01.05.02.02 | 02
vt. 02.02.20.08
vt. 03.10.06.03.06 | 09.02

temper
n. 01.01.11.02 | 09
n. 01.02.05.05
n. 01.02.08.01.22.03 | 09
n. 01.02.08.03.04.01 | 02
n. 01.04.03.03
n. 01.05.01.05.02
n. 01.05.05.21.04.03
n. 01.06.01.11.02 | 01
n. 01.06.03.01.01 | 05.02
n. 01.06.03.01.03 | 12.05
n. 01.06.07.05
n. 02.01.04
n. 02.01.04.01
n. 02.01.15.07.02 | 01
n. 02.02.06
n. 02.02.18
n. 02.02.18.02
n. 02.02.21.05
n. 03.10.12.02.03 | 10
n. 03.10.12.03.12.07 | 02
n. 03.10.12.03.12.08 | 09
n. 03.10.12.03.23 | 06
vi. 01.05.05.02.03.01 | 03.02
vi. 01.05.05.22.01.01 | 03
vi. 02.05.04.01.01
vi. 03.10.12.03.12.02 | 06
vt. 01.02.01.03 | 00; 08
vt. 01.02.08.03.04.01 | 02
vt. 01.02.08.03.08.01 | 04.01
vt. 01.04.04.01.03 | 01
vt. 01.04.05.01.01 | 06
vt. 01.04.05.04.06
vt. 01.05.02 | 06
vt. 01.05.05.20.04
vt. 01.05.05.21.04.03
vt. 01.05.06.06.01 | 01
vt. 01.05.06.11.02.01
vt. 01.06.03.01.03 | 11
vt. 01.06.06.09 | 11
vt. 01.06.07.05.04 | 00; 09; 15; 18; 19.03
vt. 02.01.15.06.01.03
vt. 02.01.15.07.08.03 | 09
vt. 02.05.06.01
vt. 03.01.07.06.01 | 03
vt. 03.04.03
vt. 03.04.09.05 | 03
vt. 03.10.06.02 | 03.01
vt. 03.10.06.03.06 | 06; 06.05; 09.04
vt. 03.11.03.01.01.03.01 | 00; 01; 04
vt. 03.11.03.02.04.06 | 01

</div>

01.06.01.09 (vi.) *Be similar*
01 *resemble/take after* braid of c1205–1691; 1864 · anliken 1340(2) · liken a1400/50; 1809–1838 · semble c1400–a1450 · draw after c1450–c1475 · show 1578–1877 (*arch.*) · assimilate (with) 1768–1851 **01.01** *be like to* resemble 1429–1836 **02** *share the nature of* participate 1531–1715 · have common (with) 1577–1608/11 · partake c1585–1620 · have of common (with) 1657 · have in common (with) 1705– **03** *be a pair/like each other* resemble 1751–a1871 · be a pair 1840– (*colloq.*) **04** *border upon/approach/approximate* coast 1382–1713 · approach c1400– · want little/nothing c1420–1784 · sympathize (with) 1591–1668 · neighbour 1631 · trench on/upon 1635–1876 · bear 1682–1835 · border on/upon a1694–1839 · approximate (to) 1789/96– · verge 1825– · begin 1833– (*colloq., orig. US*) **04.01** *in look/appearance* blush on c1530 · verge 1815– **04.02** *narrowly avoid* come near 1878– **05** *have a touch/tinge/suggestion of* sound in/into/to/towards c1340–a1661 · smatch c1380–1613 · sound of 1393; a1548 · smell 1526– · savour of 1548– · taste 1559–1840 · smack 1595– · relish 1602– · season of 1618 · reflect 1653 · tincture 1787 · twang 1821 **05.01** *taste/savour of* taste 1596 (*Scots*) **06** *become like* anliken 1340(2) · liken a1400/50; 1809–1838 · assimilate to/with 1768–1851 **07** *correspond* run together a1225– · answer c1230– · agree (to/with) 1526– · correspond c1645– · analogize 1733– · homologize 1733– · begin 1862– (*US*)

THE IMPORTANCE OF THE WORK

- 'The OED gave us individual trees, but never a sight of the whole forest or helpful pathways through it. The thesaurus does precisely that.' *David Crystal*

- **'Forty-five years of exacting scholarship by a well-led team have had a triumphal outcome.**

- This book is a magnificent achievement of quite extraordinary value. It is perhaps the single most significant tool ever devised for investigating semantic, social, and intellectual history.' *Randolph Quirk*

Reception

- ***Historical Thesaurus* is a masterpiece worth waiting 40 years for (Guardian)**
- **After a 44-year labour of love, world's biggest thesaurus is born (Daily Telegraph)**
- **The largest thesaurus in the world, and the first historical thesaurus in any language.**
- **I can hardly imagine any reference book more valuable.** *Philip Pullman*

0. TLL, OLD, CGD HTOED : **solid state** databases

- By contrast to HTOED TLS invites and is and has been for many years something of a WIKI research discussion platform for students of classical Chinese:

- TLS involves and enables continuous expansion, correction and developement at all levels of its complex organisation.

- TLS will enable adjustment and expansion of

- a. attributions of word meanings,

- b. syntactic functions of words,

- c. semantic features of words,

- d. definitions of words, of their functions, and of their semantic features,

- e. definition and attribution to text passages of rhetorical devices,

- f. synonym discrimination for words and lexicalised expressions. (see also 成語同義詞典

Paper!

- g.systematic analysis of the way the meanings of a word derive from/relate to its root meaning.

- h.explicit morphology

Digitised sources we have, but not yet integrated into TLS

- 110 000 inscriptions from 1300-800 BC, transcribed
- 14 000 bronze inscriptions 1200 -700 BC, transcribed
- 67 000 words defined in modern Peking Mandarin Audio Idiolet Dictionary
- 10 000 proverbs defined in Peking Mandarin Audio Idiolect Dictionary
- http://tekstlab.uio.no/maid/
- http://maid.nowa-huta.krakow.pl/

- A very large number of outside vast corpora available
- https://ctext.org/
- http://ccl.pku.edu.cn:8080/ccl_corpus/
- which may be usefully compared with what we have.

Missing sources for analysis we have, but not yet integrated into TLS:

- 說文解字

- !

- Daozang

HTOED is based on dictionaries

- Based mainly on *Oxford English Dictionary*
- Based also on *Old English Wordlist*
- Covers 1000 years.
- TLS consults a digitised set of standard dictionaries but consists of independent reading of original texts.
- TLS is integrated into a comprehensive primary source database.
- TLS can base itself on contrastive historical dictionaries of a type that do not exist for English:
- Yuan Hui, "Contrastive Dictionary of Ancient and Modern Chinese" 1561 pages, Taiyuan, 1994
- Many such dictionaries are being used.

1. TLL HTOED has no **systematic** morphology

- All lexical entries that consist of more than one character are to be provided with a morphological analysis in TLS:

- kexue 科學 is not just glossed as "science", but the lexical entries in TLS for 科 and 學 relevant to the interpretation of this word are to be identitied.

- Generally, in FileMaker TLS every binome XY is to be analysed by explaining exactly which of the many meanings of X, and which of the many meanings of Y combine in exactly which syntactic constructions to form the complex word XY.

- In FileMaker TLS the English word *disingenious* would be related to the dictionary item for *dis-*, *in-*, *geni-* and –ous, as I do suggest any linguistically adequate dictionary of English would have to do.

- Morphological and syntactic functions of Chinese characters may not always be the same: **which of the 25 *ab* and *ab-* in Oxford Latin Dictionary 1982 do we have in *abusio* "abuse"? These are the questions TLS aims to answer precisely.**

2. HTOED is a word-thesaurus **not** a conceptual thesaurus

- Taxonomic headwords are **undefined**.

- Synonyms are **undiscriminated**.

- TLS includes as a central part a metalinguistic taxonomic ontology of CONCEPTS that are defined in terms of other **defined CONCEPTS.**

- Being about concepts, not words, the taxonomic ontology of CONCEPTS in TLS enables and invites cross-linguistic comparison with such languages as Latin and Greek

- Source references in TLS for the metalinguistic CONCEPTS include systematic references to Greek, Latin, French, Russian, German and Carl Darling Buck, *A Dictionary of Selected Synonyms in Indo-European Languages*, Chicago University Press.

3. HTOED says nothing about the currency of words and treats highly current words along the same lines as very rare words.

- A *hapax legomenon* (word which is only recorded once in history) would be listed in HTOED alongside a common and garden household word for the same thing.

- *Thesaurus Linguae Sericae* tries to rank currency and will glosses style where appropriate. (This must be expanded.)

- TLS distinguishes between highly **current default vocabulary** (for which there is a summary synonym dictionary compiled in close cooperation with Professor Jiang Shaoyu in TLS) from **less common expressions** so far not entered into the synonym dictionary.

4. HTOED operates with six word classes

- n, vi, vt, adj, adv, prep
- *Thesaurus Linguae Sericae* distinguishes 1 200 word classes.
- For example among nouns among many others:
- 1.abstract nouns "thirst"
- 2.mass nouns "milk"
- 3.count nouns "apple"
- 4.collective nouns like "cutlery"
- 5.generic nouns like "edibles"
- 6.relational nouns like "brother"
- 7.proper nouns like "Socrates"

Syntactic Functions

- TLS distinguishes between three basic word classes (nouns, verbs, and particles) and derives from these over 1 000 logically and grammatically distinct functions of expressions in classical Chinese so far found in the literature.

- A survey of these can be found in *Clavis Serica, A Survvey of Classical Chinese Syntax*, downloadable from *https://ku-dk.academia.edu/ChristophHarbsmeier*

5. HTOED establishes no systematic lexeme relations between words.

- **Near-synonym-distinctions** are not made explicit in HTOED.
- **Antonyms** are constitutive of understanding a word and a concept: good/bad.
  - One word even in one meaning usually has **several** antonyms.
- **Converses** are not given: give/receive; offer/accept
- **Opposites** are not specified: eat/drink
- **Epithets** are not listed: **humble** servant
- *Thesaurus Linguae Sericae* systematises all these and more.

6. HTOED disregards pronunciation history

- Whether *sun* and *son* were homophones in Shakespeare's or Milton's times, or in Chaucer's times remains unrecorded in HTOED as it does in OED.

- *Thesaurus Linguae Sericae* makes such historical features systematically available.

7. HTOED disregards orthographic history

- The earliest large-scale Chinese dictionary of AD 100, known as the Shuowen, is entirely devoted to the systematic analysis of Chinese orthography.
  - There is no orthographic dictionary of English even today.
- There are, for example, 490 ways of spelling the word **through**, disregarding dialect spellings, of which only an arbitrary selection is messily laid out in OED.
- For tracing termiological history, spelling ambiguity is crucial.
- *Thesaurus Linguae Sericae* makes these systematically available.

8. HTOED disregards rhetorical devices: but these are central to conceptual history, for very good reasons

- Parenthesis,

- Irony,

-

- *Thesaurus Linguae Sericae* provides a taxonomy of over 600 such systematically defined rhetorical devices 辭格 and aims to account for their history in Chinese literature, because they are crucial for conceptual history.

9. HTOED gives no quick access to usage **in context**

- *Thesaurus Linguae Sericae* aims to provide fully annotatable context AND TRANSLATION of all primary source references.
- This is crucial for conceptual history.

10. HTOED does not provide defined taxonomies of its semantic fields and conceptual schemes

- *Thesaurus Linguae Sericae* provides such taxonomies for
  - 1. semantic fields
  - 2. syntactic categories
  - 3. rhetorical devices
- In the spirit of Leibniz in such a way that all Concepts in the ontology of TLS are defined in terms of other Concepts which themselves are defined in the system. Circularity is minimised but obviously inevitable in certain cases.

11 HTOED disregards the *egestas linguae*, i.e. conceptual shortfalls as applied to English as a whole

- For HTOED English is the standard for the description of itself.

- TLS assumes that like all other languages Chinese will have semantic gaps, that is to say Concepts in the overall ontology which are empty for Chinese.

- Compare Barbara Cassin, *Vocabulaire européen de la philosophie: Dictionnaire des intraduisibles* and its English mistranslation (Princeton University Press).

12. HTOED relies predominantly on printed published sources

- But: Edited printed language is secondary.

- For the last fifty years at least, there are plenty of very important un-meddled-with unedited primary sources available.

- Orally and especially video-recorded spontaneous sources are of much greater interest than the printed.

- *Thesaurus Linguae Sericae* will analyses systematically recorded oral material

- http://tekstlab.uio.no/maid/

- http://maid.nowa-huta.krakow.pl/

Chronology

- All text names must be preceded everywhere by the label identifying their date: never 左傳， always MiddleWarringStates/ 左傳 .

- All attributions must always visibly be to dated texts.

- All Concept boxes relating to a given character must have an index attached to them with the label for the period where they given character has so far been attested in TLS.

- This label must automatically change when relevantly new attributions are made.

Intertextuality

- TLS allows for the justaposition of texts that belong together, not only as text and commentary, but also as

- text and variant version of the same text

- text and other text with material relevant to the other text

- text and other contrastive text to be compared.

Cross-cultural intertextuality: desiderata

- TLS permits the juxtaposition of Greek or Latin texts with Chinese texts
- TLS will enable the application of Liddell, Scott, Jones and Lewis and Short to Latin and Greek texts within the database.
- TLS will enable the annotation of Greek and Latin texts by the creation of supplementary lexicographic entries on Greek and Latin words important to conceptual history.
- TLS will compile a limited corpus of Graeco-Roman texts for the purpose of systematic study of intertranslatability with Chinese, to begin with, Aristotle, Plato, Marcus Aurelius, Horace, Catullus, St. Augustine, Thomas Aquinas.

Sources

- All source references must be marked for presence or absence and OCR status of pdf.
- References to Sources must be clickable so that the relevant page comes up for those with the right advanced pasword.

Currency rating among the SW of a word

- Frequency of usage (not only of words) must be indicated
- default
- dominant
- frequent
- current
- limited
- rare
- hapax

Frequency

- Any character looked up in TLS will have a number indicating its relative frequency in the corpus so the student can assess how worth while it is to learn the character in question.

Style

- A range of standardised stylistic glosses may be added for non-ordinary word meanings at the start of definitions:

- poetic

- vulgar

- colloquial

- technical term

- etc

TLS is Work in Progress: Every detail is open to change, correction, **and deletion**

- Any field in which change is possible with passwords should have a green frame to show change is possible for one.

1988-	Production, collection and revision of digitalised classical Chinese texts as a raw annotated database searchable only in Nisus.
1993-	Production (with much enthusiastic student help) of **interlinear bilingual editions** of pre-Buddhist Chinese texts for personal use as a way of looking over the shoulders of and learning from the best translator-sinologists as an information database. The **University of Oslo** provides limited but crucial and sustained financial support for this digitilisation project over many years.

1994-

Compilation of a **comprehensive synonym dictionary** of classical Chinese, produced as a complex formatted word-processing document for the convenience of students and as a personal record as an analytic database.

1997-

Jens Østergaard Petersen (Copenhagen) begins to to develop a **FileMaker database** to accommodate the increasing technological needs of the developing database.

1998-

TLS becomes a **cooperative project** between Peking University, the University of Heidelberg the University of Oslo and a large number of scholars from other Universities from whom TLS has received generous help.

1999-2000	TLS is supported by the Institute for Advanced Study, Oslo, and one year's funding at Norwegian professorial level is made available to J.O. Petersen.
2001	The Chiang Ching Kuo Foundation finances the **Web-publication** of parts of TLS to be published in the Documentation Centre of the Sinological Institute, University of Heidelberg. A $38 000 grant is made available to be used in its entirety to support the programming work of J.O. Petersen under the administration of Prof. R. G. Wagner, Dept of Chinese, at the University of Heidelberg.
2004-2015	TLS is supported by the **Swedish Institute for Advanced Study** in the Social Sciences at Uppsala.

Comparability

- TLS assumes that all cognitive cultures are in some important sense **not exactly translatable or even commensurable**, and that this is what makes the close philological study of languages profoundly rewarding and necessary.

- Through globally defined SYNONYM GROUPS, TLS aims to **cultivate a sense for the specificities of Chinese cognitive culture by making this culture explicitly comparable** in every possible detail to that of other ancient civilisations like Greek, Roman, and Hebrew etc,

- and especially also with languages like Russian, which is the language on which the most detailed work on synonymy has been done by Juri Apresjan and his many associates.

- The medium of the TLS cross-cultural comparative analysis through SYNONYM GROUPS has to be a **maximally universal metalanguage**, preferably one which is based on
1. detailed philological experience with **a variety of "classical" and other languages where synonymy relations have been studied in detail.**
2. a systematic application of **logical, grammatical, and semantic analysis.**

Original features in TLS

- TLS is the first dictionary which systematically **organises the Chinese vocabulary in taxonomic ("is a kind of") and mereonomic ("is part of") hierarchies**, thus showing up whole conceptual schemes or cognitive systems. These are taken to circumscribe the changing topology of Chinese conceptual space.

- TLS is the first dictionary that **systematically registers a range of lexeme relations**, such as antonym, converse, epithet, etc. TLS thus aims to define conceptual space as a relational space, consisting of

- TLS is the first dictionary of Chinese which incorporates **detailed syntactic analysis of (over 900) distinct kinds of) syntactic usage.** TLS thus enables us to make a systematic study of such basic phenomena as the natural history of

- For example, TLS therefore also makes possible the systematic study of such basic phenomena as the history of **abstraction and reflexivity** in Chinese.

- TLS is the first corpus-based dictionary which will record **the history of rhetorical devices** in texts and will thus enable us to study such intellectually crucial things as the natural history of irony in China.

- TLS includes the first dictionary of **290 mathematical terms**, by Karine Chemla, CNRS, Paris.

- TLS is the first dictionary of classical Chinese aiming to be systematically **linked to archaeological illustration** of whatever can be illustrated from excavated evidence. (For an English index to Käthe Finsterbusch's magisterial motif index to Han Dynasty mural illustrations, search for

Current state

- TLS is designed not as an authoritative reference work but as **a digital construction site, a convenient platform for the continuing discussion** on the development of the Chinese language over the last 3000 years.

- All distinctions and contrasts proposed in TLS aim to state no more than **preliminary but explicit and refutable working hypotheses** on Chinese historical grammar, concepts, and rhetoric.

- TLS is still at the initial stage of trying to build up a comprehensive and convenient **philological infrastructure** that will enable one to discuss an increasingly sensitive analysis of the the lexicon, the grammar, and the rhetoric of the Chinese language in its historical development over the last 3000 years.

- TLS is, and will try to remain, cheerfully over-ambitious, **exploratory and experimental**.

- TLS started out and still largely remains the chief editor's **only partially systematised collection of research notes** concerning analytical work in progress on the history of Chinese conceptual schemes. TLS still retains many traces of its fumbling humble origins throughout, in spite of the very **substantial contributions from many distinguished collaborators.**

- TLS is not designed as an exercise in formal semantics or logical analysis, but as an **accessible and convenient and pedagogical digital vademecum for fellow students** who, like we, are interested in the historical development of Chinese cognitive schemes, Chinese words, and Chinese grammatical structures.

Autochthonous perspectives on synonymy

- **An Autochthonous Perspective**

- TLS insists on looking upon matters Chinese from a Chinese perspective, and the working language of the project is for a large part Chinese. Ca. 700 Chinese books and articles on Chinese synonymy have been listed and assembled/photocopied so far, for systematic analysis. This represents a rich Chinese tradition of scholarship.
Representative examples of Chinese contributions towards conceptual history include the following:
Wang Fengyang, "Dictionary of Classical Chinese Synonyms", 1053 pages, Changchun 1998
Huang Jingui, "Classified Contrastive Dictionary of Ancient Chinese Cultural Vocabulary" 1594 pages, Shanghai 1995
Chen Bingzhao et al.,"Detailed Explanatory Dictionary of (Modern Chinese) Synonyms and Antonyms", 1000 pages, Changsha, 1996
Dong Danian, "Classified Dictionary of Contemporary Chinese", 1334 pages, Shanghai 1998
Jiang Guanghui, "Modern Chinese Current Use Contrastive Dictionary", 944 pages, Changchun, 1998

Conceptual Schemes and Synonyms

- TLS explores the **conceptual schemes** of pre-Buddhist Chinese on the basis of a **corpus of translated texts** interlinked with an **analytic dictionary**.

- Text and dictionary are constantly held up against one another. Our understanding of the texts and the Chinese system of meanings can be refined by through this close confrontation.

- TLS associates Chinese concepts with concepts from the European antiquity, aiming to make the classical Chinese evidence comparable to that of other cultures.

- TLS seeks to make more precise the criteria used in translating classical Chinese, through a detailed description of the semantic relations that obtain among Chinese words.

- TLS is the first **synonym dictionary** of classical Chinese in any Western language; it attempts to state as clearly as possible the semantic nuances that distinguish words close in meaning.

- TLS is the first dictionary which systematically organises the Chinese vocabulary in **taxonomic and mereonomic hierarchies**, thus exploring the topology of the Chinese mental space.

Globalised Conceptual Ethnography

- TLS received much practical help from such **synonym handbooks** as
  L. Doederlein, Lateinische Synonyme und Etymologien. 6 Bde (Leipzig: Vogel, 1826-36)
  J.H.H. Schmidt, Handbuch der lateinischen und griechischen Synonymik (3 vols. Leipzig: Teubner, 1889)

  TLS aims for grammatically explicit and philologically **detailed semantic portraits of the meanings of keywords** in the spirit of Juri Apresjan and Igor A. Mel'cuk:
  Juri Apresjan, A New Explanatory Dictionary of Russian Synonyms (1450 pages, 2005)
  Juri Apresjan, Systematic Lexicography (Oxford: OUP, 2001)

  I.A. Mel'cuk and Alexander K. Zholkovsky, Explanatory Combinatorial Dictionary of Modern Russian, (in Russian)
  Vienna: Slawisischer Almanach, 1984

  TLS has profited from the courageous attempts at **contrastive conceptual ethnography by Anna Wierzbicka**:

- Semantics, Culture, and Cognition: universal human concepts in culture-specific configurations, OUP 1992
- Understanding Cultures Through Their Key Words, OUP 1997
- Emotions across Languages and Cultures, Cambridge Univ. Press 1999

Conceptual History

- TLS was inspired from the start by the tradition of **conceptual history** as represented in the work of the social historian R. Koselleck, and as exemplified in the comprehensive, though Euro-centric

- *Geschichtliche Grundbegriffe [Basic historical concepts] 8 vols.*

- *Historisches Worterbuch der Philosophie [Historical dict. of philosophy] 12 vols*

- *Historisches Worterbuch der Rhetorik [Historical dict.of rhetoric] 6 vols,*

- *Archiv fur Begriffsgeschichte [Archive for conceptual hist.]40 volumes so far. (1955-))*

- TLS sees the history of conceptualisations and of conceptual schemes as partly creating and defining **the evolving intellectual and social space in which historical agents think, plan, and operate.** (See the seminal R. Koselleck, *Kritik und Krise* (1959), tr. MIT Press 1988).

  TLS aims to provide a **bilingual, broad philological basis** for the comparative investigation of such conceptual developments in China.

  The main historical focus of TLS is on the **operative common-use current vocabulary** which shaped general Chinese cultural and social practice over the last 3000 years, with special emphasis on **keywords** of the civilisation. (See R. Williams, *Keywords* (Fontana, 1985))

Lexical Relations

- TLS is the first dictionary that systematically registers **lexical relations** like antonym, converse, epithet, etc., thereby aiming to define the Chinese conceptual space as a relational space.

- TLS is the first dictionary of Chinese which incorporates detailed **syntactic analysis**, thus enabling the systematic study of basic phenomena as e.g. the history of abstract nouns in China.

- TLS is the first corpus-based dictionary which records the history of **rhetorical devices** in texts, making it possible to study crucial matters such as the history of irony in China.

- **Navigate the TLS site** by clicking the buttons above. The box on the **left** tells you where you are. The boxes on the **right** tell you where you can go and function as buttons. If the labels on the buttons have ">" in front of them, the button has a pulldown menu; if you click on it instead of using the pulldown menu, it functions as an ordinary button.

- The **fastest way** to learn what TLS has to offer is by using "Basic Search". Click here to choose Basic Search. In Basic Search, choose whether you wish to search for a Chinese or an English expression and then what you wish to search for – a word, a synonym, a passage in a text or a translation, etc.

- Another **shortcut** is to access the list of TLS "File List", by first clicking here and then clicking on a button in the "Sample Records" column. A Sample Record is an example of a full TLS record showing the wide range of information contained within TLS. – You may also click on "Random Records" on the File List, but you may then end up in a place that is not so informative (or not so well finshed by us).

Oracle Bone Inscriptions and Bronze Inscriptions

- TLS includes one year's work of careful analysis of Oracle Bone Inscriptions.

- The transcriptions 釋文 of all OBI in the many volumes of the standard 合集 are searchable, and high-quality rubbings based not on 合集 but the books 合集 was photocopying form are available in TLS, so that the history of the language can be traced to its earliest recorded phases.

- Bronze inscriptions are ready to be annotated within the system.

- For the study of the early history of prehistory of concepts these sources are obviously crucial, and they are available for translation annotation in TLS.

- Leading scholars have worked on Oracle Bone Inscriptions for one year.

Chronology

- Pre-Classical Archaic Chinese, which is studied on the basis of oracle bones (presented and analysed by Ken-ichi Takashima) and bronze inscriptions (presented and analysed by Ulrich Unger).
- Pre-Buddhist Classical Chinese, which is studied on the basis of the epigraphic and traditional written sources down to the second century A.D. which have remained influential throughout the ages, but for which there is a large amount of newly discovered excavated textual evidence.
- Medieval Colloquial Chinese, for which some of our most important sources are translations from the Sanskrit and other Buddhist colloquial writings. For this period it becomes crucial to take account of Buddhist influence on Chinese grammar and lexicon.
- Modern Peking Chinese, for which we do at last possess the kind of primary oral evidence which was lacking for the preceding stages of the language. For this period it becomes crucial to take account of the overwhelming western influence on Chinese grammar and lexicon, and to make use of the availability of oral/filmed modern evidence.

Diachrony

- **Cross-Cultural Diachrony**

- TLS aims to describe the long-term Chinese conceptual history in a **dynamic relation to modern developments**: it concentrates first on classical Chinese and will then relate this to Modern Chinese by a **rigidly homogeneous** analysis of both along the same principles .
  Speakers of Modern Chinese tend to have an active vocabulary of **thousands of phrases with classical Chinese lexis and grammar**. Early colloquial elements are even more prominent. TLS will register these continuities as **the current presence of the past** in Chinese.
  A large part of **modern Chinese morphology is explained by the principles of classical syntax**. The morphemes of many Modern Chinese words will be explained by reference to the relevant TLS lexeme entries for the classical Chinese words from which they derive.
  TLS will register **Buddhist, Mongol,** and more recent **westernising influences** on the formation of the core Modern Chinese vocabulary.
  TLS distinguishes between terminologised and general vocabulary, and it will **concentrate on the historical and contrastive analysis of CURRENT Chinese core vocabulary and key terminology**.

Ranking of Meanings

- **The Ranking of Meanings**

- Lexical entries (meanings) in TLS are to be ranked with respect to their currency on a scale from 5 "extremely common default meaning" to 1 "hapax legomenon, i.e. attested only once (or twice)".
  According to the definitions in Collins English Dictionary a speaker wishing to understand the sentence "The cat sat on the mat" should have to choose between 41 310 000 combinations of meanings of theses words in order to understand the sentence, given that the number of meanings in that dictionary is cat (24), sit (18), on (25), mat (17), the (15). (J. Sinclair 1998)

- Clearly, this is NOT the kind of choice a speaker of English makes. Part of the reason is that word meanings are ranked with respect to currency in the native speaker's mind. TLS aims to rank lexical entries for degree of currency gradually, as a sufficiently varied corpus of analysed data becomes available. The number of cases in TLS registered does NOT, of course, necessarily represent degree of currency, and therefore this assignment of rank will remain subjective.

Grammar

- TLS aims to provide a **systematically unified account of the lexicon and of grammar**, and also of rhetoric.

- This means that **the TLS dictionary is systematically grammatical in focus**, and that the **TLS grammar is systematically linked to the dictionary**, and that **rhetoric is integrated into the analysis of texts**.

- Any semantic or syntactic category in the TLS dictionary is taken to be significant only to the extent that it is systematically elucidated in the grammar in such a way that **the problems connected with assigning the relevant category in concrete cases are made explicit**.

- Any rule or category in the TLS dictionary is taken to be significant only to the extent that its illustration through translated examples is systematically accessible in the dictionary so that **its adequacy and its limitations can thus be tested on the translated texts in TLS**.

- The relevance of TLS for modern Chinese political history, and for the globalisation of current conceptual history (Begriffsgeschichte), is discussed in:
  Christoph Harbsmeier:
  Concepts that make multiple modernities: The conceptual modernisations of China in historical and critical perspective.
  [pdf].

- The analysis of Chinese rhetoric forms an important part of TLS. A detailed discussion of the issues involved in such an analysis is found in this book-length treatment:
  Christoph Harbsmeier: *Chinese Rhetoric in Comparative Perspective* (2002) [.pdf - LARGE FILE! 11MB].

Indices

- TLS provides the means to create indices not only of the complete material in the database but also of specified books or chapters from separate books only,

Illustrations and Archaeology (Finsterbusch)

- Many of the culturally important and historically influential Chinese conceptual schemes can only be studied meaningfully if **systematic attention is paid to the fast-growing archeological evidence concerning ancient Chinese material culture.**

  TLS aims to illustrate everything that may be illustrated with **ILLUSTRATIONS based - in the first instance - on archaeological material from pre-Buddhist times.** 197 synonym groups are systematically linked to archaeological illustrations.

  TLS aims to **link all discussion of material objects to standard archaeological LITERATURE** listed in the bibliography.

  TLS aims to illustrate also such **social activities** as "feasting", "acrobatics", and even "sitting", **wherever the relevant Chinese words can be associated with confidence to archaeological representations**.

  TLS will concentrate entirely on those aspects of archaeological finds **for which the contemporary ancient terminology is reasonably well-established** and included in TLS.

- **Historical phonetics:** Jingdian shiwen 7th cent. AD, Guangyun 1008 AD, and the reconstruction of Old Chinese by Pan Wuyun, Shanghai Normal University, are integrated into the phonological component of TLS. Phonological Reconstruction: 18 055 words/readings are reconstructed
- **Character Lookup:** 71 624 characters may be looked up in TLS
- **Lexical analysis:** Over 39 000 meanings of words
- **Example sentences:** Over 112 000 examples analysed
- **Syntactic analysis:** Over 1100 function classes for words defined and exemplified
- **Synonym groups:** Over 2700 synonym groups
- **Synonym distinctions:** over 1300 groups contrastively analysed

Modern Peking Mandarin

- **The Analysis of Modern Peking Chinese**
- TLS will analyse modern spoken Chinese on the basis of **recorded spontaneous speech and oral texts** and on the basis of a fine phonetic transcription of relevant parts of these, **NOT** only - and certainly not primarily - on the basis of written (even less published!) text.

  TLS will in the first instance concentrate on **idiolects**, the language of one speaker, not on a variety of speakers of Peking Standard Chinese whose speech may differ widely. The choice of idiolects is inevitably arbitrary.

  The **example material** will in the first instance be in the form of recorded sentences from the following sources:

- A 1500 hour-long **oral dictionary** provided by Tang Yunling and recorded by Oldrich Svarny of the Academy of Sciences in Prague will be the main source.
- Digitalised **and transcribed** recordings of xiangsheng (**comic dialogues**) of Hou Baolin (Peking) and Ma Sanli (Tianjin) from the sixties and seventies.
- Modern Chinese will be studied in its pragmatic action context, and for this purpose extensive use will be made of **digitialised films** with speakers of Peking Chinese, including xiangsheng and other materials.

Innovative Features on Modern Chinese TLS

- **Innovative Features of the Planned Modern TLS**

- Up until 2003, TLS has largely concentrated on pre-Buddhist Chinese. Modern Chinese will be analysed in its historical perspective. Modern morphemes will be systematically related to ancient words.

  Modern Chinese will be analysed so as to make it pervasively and rigidly comparable in every detail to the relevant earlier stages of the language.

  Exactly the same analytic framework will be adapted and applied to all periods.

  The **repertoires of classical and tentative modern Chinese synonyms have been juxtaposed** to allow for systematic historical comparison.

  Foreign influences of all kinds on modern Chinese will be systematically recorded, Westernised grammar and lexis as well as the dominant Japanese influence on modern Chinese will be recorded in detail.

  For all multisyllabic Modern Chinese words the internal morpho-syntactic structure and semantic composition will be defined.

  Modern Chinese will be treated systematically as a stress prominent language focussing on minimal pairs like **pro**gress versus pro**gress**. Two-peak words like **re-do** do again will also be identified.

Audio Dictionary of Peking Chinese

- **Modern Peking Chinese Sound Materials**

- Ca 2000 hours of tapes in the form of an audio dictionary of Modern Standard Peking Chinese by Tang Yunling, a Manchu speaker of Peking Chinese recorded in 1966-69,

- the audio material will be **analysed/visualised** with the standard Praat system so that stress patterns of words visually demonstrated,

- covering all the material in *Xiandai Hanyu pinyin cihui* "VOCABULARY OF MODERN CHINESE" (1963),

- in Lu Zhiwei, *Beijing hua danyinci cihui* "MONOSYLLABIC VOCABULARY OF PEKING SPEECH" (1956)

- and in Zhang Xunru, *Beijing hua qingsheng cihui* QINGSHENG WORDS IN PEKING SPEECH (1957).

Modern Morphology and Syntax

Analysis will be based on C. Harbsmeier, Modern Chinese Analytic Syntax, 2 vols. Oslo: Dept of East Asian Studies, 1992.

Morphological and syntactic constituent structure will be analysed according to the same principles in TLS, wherever possible.

This is possible because so much of Modern Chinese (MC) morphology is classical Chinese syntax. (W. Faulkner: The past is not dead. It is not even past!)

This is desirable in order to show how much of the lexis and grammar of classical Chinese remains an integrated core part of MC.

The meanings and syntactic functions of MC morphemes in MC word formation are registered and analysed exactly along the same principles as the meanings of words.

E.g. 𝆄 "school head" is taken as evidence that can be adnominal and mean "school", (cf. 𝆄 "campus") and that may function as a nominal head meaning "head (of an institution)" (cf. | "head of the institute").

Bound morphemes will be treated like (not as!) bound words

The complex questions of (degrees of) wordhood in MC are not allowed to get in the way of a homogeneous and strictly systematic analysis of constituent structure in MC.

Levels of Analysis for Modern Spoken Chinese

- TLS focusses on the following:

  Modern Chinese will be treated as a stress prominent language and stress assignment will be verified by the study of visualised analysis of the phonetic formants.

- **the modern conceptual repertoires** and systematic lexical relations rather than simply on word meanings.

- The informants spontaneous **subjective reconstruction** of the meanings of words in her repertoires.

- The informants **actual linguistic practice** (which may exemplify features which are not included in her subjective reconstruction).

- The **contrasts between the informants subjective reconstruction and the normative definitions** in the standard Chinese-Chinese dictionaries and grammars of the time.

- The contrasts between her **practice** and that described in the standard Modern Chinese **dictionaries and grammars** of the time.

- The contrasts between the language of our **acoustic material** on the one hand, and the contemporary **published sources** on the other.

- The contrasts between the published Modern Standard Chinese **material** and the standard **description** of this language in the standard dictionaries.

Rhetorical Devices

- Definitions cannot be changed must be made writable in.
- New devices cannot be entered, must be added.
- New attributions cannot be entered, must be enabled.

- Attributions get mixed up with the lexical taggings

Source references

- These must be made clickable whenever we have pdfs available throughout the system.

Texts: Titles

- The More Texts section is disorganised.
- When there are several versions of the same text these have to come up together.