

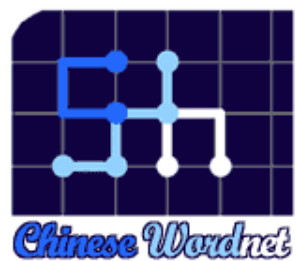
Language Resources and Sense- Related Research: An Overview

現代漢語語言資源總覽

Chu-Ren Huang

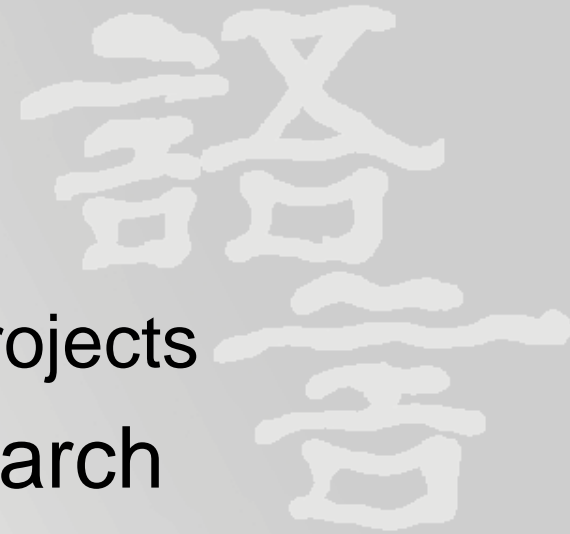
Institute of Linguistics, Academia Sinica

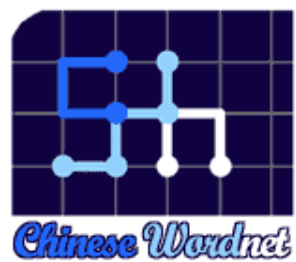
<http://cwn.ling.sinica.edu.tw/huang/huang.htm>



Outline

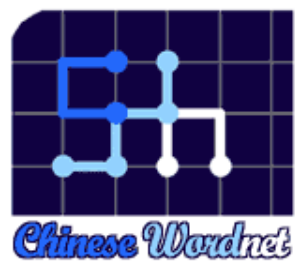
- Background:
 - Historical Milestones
 - Current Research Issues and Projects
- Language Resources as Research Infrastructure
 - Corpora and Extension
 - WordNet/Ontology





Background: Historical Milestones

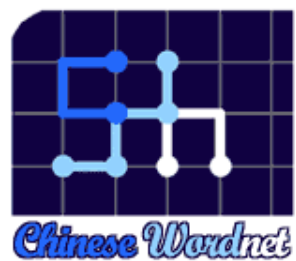
- **1986:** CKIP (Chinese Knowledge Information Processing) Group and First Chinese Lexical KnowledgeBase Project (Ching-chun Hsieh and Keh-Jiann Chen, Funded by ITRI)
- **1990:** First Chinese Corpus Project: Linguistic Analysis of Computerized Classical and Modern Chinese Text-Corpora- A Pilot Study in Chinese Corpus Linguistics (Chu-Ren Huang and Paul Thompson, funded by CCK Foundation)



Historical Milestones II

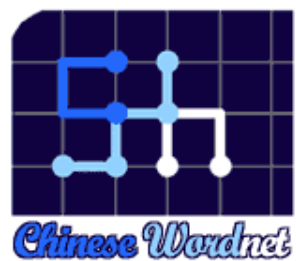
- **1989:** CKIP Electronic Dictionary
- **1992:** Modern Chinese Corpus (1 million words)
- **1993:** Pre-Qin Corpus (P.C. Wei, Thompson, Huang, Chen)
- **1994:** Sinica Corpus 1.0: First balanced and tagged Mandarin Chinese Corpus
- **1995:** CKIP Lexicon with Grammatical Information

PI's for resources on this page are Chen and Huang, unless noted otherwise.



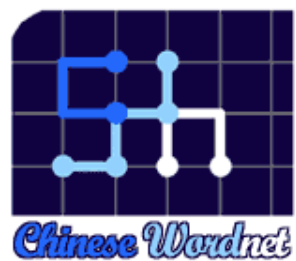
Historical Milestones III

- **1996:** Sinica Corpus on WWW (November)
 - Current Website:
<http://www.sinica.edu.tw/SinicaCorpus/> (Chen and Huang)
- **1999:** CNS 14366 National Standard for Chinese Segmentation (Huang)
- **1998-2000:** Digital Library and Digital Museum Projects
 - SouWenJieZi <http://words.sinica.edu.tw/> (Huang)
 - Adventures in Wen-Land <http://www.sinica.edu.tw/wen/> (Huang)
 - Litchi and Mirror- Classical Southern Min (Feng-Ju Lo)
http://cls.hs.yzu.edu.tw/lm/Lm_home.asp



Historical Milestones IV

- **2001:** 'Digitization of National Archives' project included Language Archives as a sub-branch (PI: Huang)
- **2002:** Both Language Archives and the spin-off Linguistic Anchors became sub-projects of a the National Digital Archives Project (NDAP).



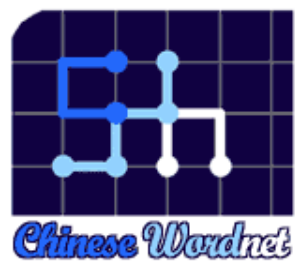
Background: Current Research Issues and Projects

- From Context to Meaning (and Back)
 - Linguistic definition and identification of sense in Mandarin Chinese (supported with corpus data)
 - Linguistic definition and identification of grammatical relations in Mandarin Chinese
 - Automatic tagging and prediction of senses
 - Account for meaning extension in context
 - Meaning facets
 - Paronymy
 - metaphor



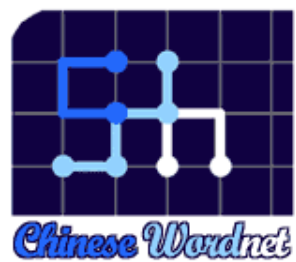
Background: Current Research Issues and Projects

- From Concept to Meaning (and Back)
 - OntoLex: The Interface between Ontology and Lexicon
 - Discovery of conceptual structure through linguistic Convention
 - From Swadesh List to Skeletal Ontology
 - From Chinese orthographic convention to linguistically encoded domain ontology
 - Cross-domain and Cross-lingual Representation of meaning and concept



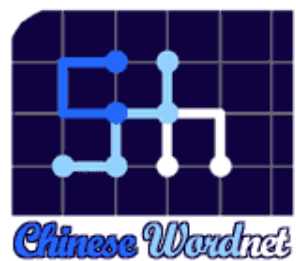
Background: Current Research Issues and Projects

- From Context to Meaning (and Back)
 - Modeling and Measurement of Meaning (M3).
2008.01-2010.12. NSC (An ANR-NSC Bilateral Collaborative Project with University of Toulouse)



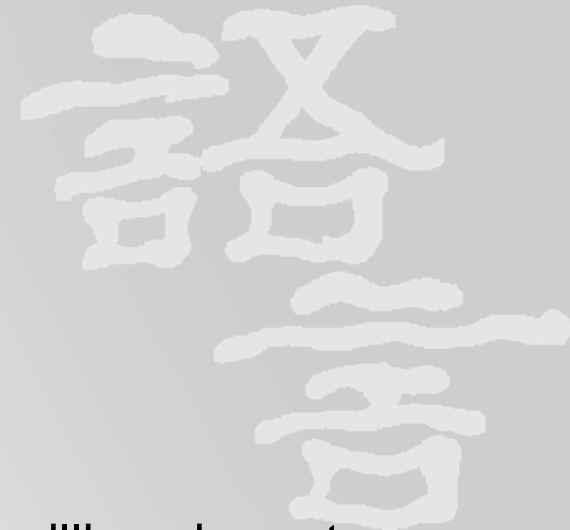
Background: Current Research Issues and Projects

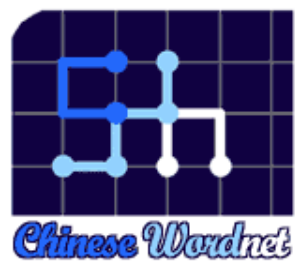
- From Concept to Meaning (and Back)
 - KYOTO: Knowledge-Yielding Ontology for Transition-based Organizations. 2008.03-2011.02. NSC (As a partner in a European Union Framework 7 project)
 - Developing International Standards of Language Resources for Semantic Web Applications. 2005.10-2008.009. NEDO Foundation, Japan. (Partner of a 5-nation group headed by Titech)
 - ISO standard for Lexical Representation (so-call LMF)
 - http://en.wikipedia.org/wiki/Lexical_Markup_Framework
 - Lexicon-driven Ontology and Conceptual Structure. 2005-2009. Academia Sinica Investigator Project



Language Resources as Research Infrastructure

- Corpora and Extension
 - From Context to Meaning
 - Sinica Corpus
 - Chinese Word Sketch
 - Tagged Chinese Gigaword Corpus (1,400 million characters, published by LDC, tagged by Academia Sinica)
 - Taiwan Presidential Corpus





Chinese WordSketch: An Overview

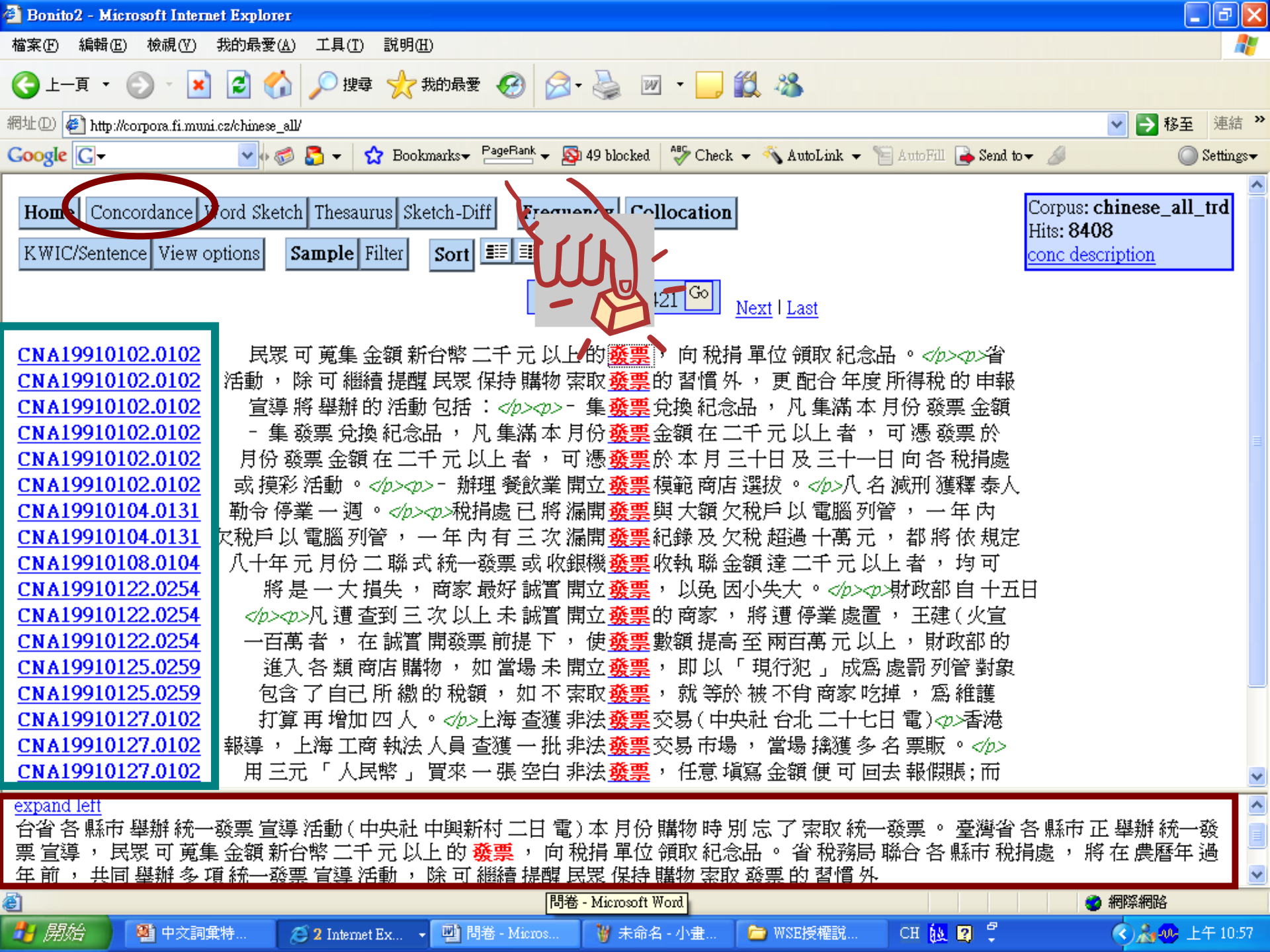
<http://wordsketch.ling.sinica.edu.tw>

- Concordance
- WordSketch
- Sketch Difference
- Thesaurus

Word Sketch Engine

<http://www.sketchengine.co.uk>

NOTE: Chinese Word Sketch can be accessed free of charge in Taiwan through our collaboration with Lexical Computing Ltd. Accessing data in any other language or outside of Taiwan must be licensed directly from Lexical Computing Ltd.



Home **Concordance** Word Sketch Thesaurus Sketch-Diff Frequency Collocation

KWIC/Sentence View options Sample Filter Sort [List Icon] [List Icon]

Corpus: chinese_all_trd
Hits: 8408
[conc description](#)

121 Go Next | Last

- [CNA19910102.0102](#)
- [CNA19910102.0102](#)
- [CNA19910102.0102](#)
- [CNA19910102.0102](#)
- [CNA19910102.0102](#)
- [CNA19910102.0102](#)
- [CNA19910104.0131](#)
- [CNA19910104.0131](#)
- [CNA19910108.0104](#)
- [CNA19910122.0254](#)
- [CNA19910122.0254](#)
- [CNA19910122.0254](#)
- [CNA19910125.0259](#)
- [CNA19910125.0259](#)
- [CNA19910127.0102](#)
- [CNA19910127.0102](#)
- [CNA19910127.0102](#)

民眾可蒐集金額新台幣二千元以上的**發票**，向稅捐單位領取紀念品。省活動，除可繼續提醒民眾保持購物索取**發票**的習慣外，更配合年度所得稅的申報宣導將舉辦的活動包括：

- 集**發票**兌換紀念品，凡集滿本月份**發票**金額
- 集**發票**兌換紀念品，凡集滿本月份**發票**金額在二千元以上者，可憑**發票**於月份**發票**金額在二千元以上者，可憑**發票**於本月三十日及三十一日向各稅捐處或摸彩活動。
- 辦理餐飲業開立**發票**模範商店選拔。

八名減刑獲釋秦人勒令停業一週。稅捐處已將漏開**發票**與大額欠稅戶以電腦列管，一年內欠稅戶以電腦列管，一年內有三次漏開**發票**紀錄及欠稅超過十萬元，都將依規定八十年元月份二聯式統一發票或收銀機**發票**收執聯金額達二千元以上者，均可將是一大損失，商家最好誠實開立**發票**，以免因小失大。財政部自十五日

凡遭查到三次以上未誠實開立**發票**的商家，將遭停業處置，王建(火宣一百萬者，在誠實開發票前提下，使**發票**數額提高至兩百萬元以上，財政部的進入各類商店購物，如當場未開立**發票**，即以「現行犯」成為處罰列管對象包含了自己所繳的稅額，如不索取**發票**，就等於被不肖商家吃掉，為維護打算再增加四人。

上海查獲非法**發票**交易(中央社台北二十七日電)香港報導，上海工商執法人員查獲一批非法**發票**交易市場，當場擒獲多名票販。用三元「人民幣」買來一張空白非法**發票**，任意填寫金額便可回去報假賬；而

expand left
 台省各縣市舉辦統一發票宣導活動(中央社中興新村二日電)本月份購物時別忘了索取統一發票。臺灣省各縣市正舉辦統一發票宣導，民眾可蒐集金額新台幣二千元以上的**發票**，向稅捐單位領取紀念品。省稅務局聯合各縣市稅捐處，將在農曆年過年前，共同舉辦多項統一發票宣導活動，除可繼續提醒民眾保持購物索取**發票**的習慣外。

Home Concordance **Word Sketch** Thesaurus Sketch-Diff

發票 chinese_all_trd:taiwan-only freq = 8408 [change options](#)

object_of 928 5.2		subject_of 692 0.3		a_modifier 276 0.5		n_modifier 562 -12.3		modifies 607 -12.8	
開立	253 73.78	收執	15 43.41	不實	85 57.41	增值稅	103 47.76	逃漏稅	16 30.08
虛開	50 51.4	給	69 32.61	假	48 46.83	銷貨	17 34.96	存根聯	7 29.53
偽造	70 40.88	對獎	7 29.59	空白	24 41.9	收銀機	16 34.19	金額	55 27.38
虛購	12 40.47	中獎	12 25.43	中獎	22 39.32	式	32 30.07	面額	13 25.7
漏開	16 40.07	逃漏	9 24.03	普通	19 28.32	小額	23 29.64	日期	18 21.74
開具	30 39.48	充當	9 23.46	領用	5 27.62	進項	9 27.8	獎金	18 20.77
變造	14 27.67	兌換	16 23.27	可疑	5 15.34	聯	29 27.75	偷稅	7 20.53
虛設	7 22.83	換版工作	3 22.92	增值	3 12.82	虛立	4 26.96	助創世	2 19.68
使用	49 22.27	捐贈	13 22.15	原始	3 11.24	加副聯	3 22.14	號碼	12 19.14
開出	15 21.51	換好	4 21.78	填開式	1 11.17	愛心	18 21.81	影本	7 18.24
購買	24 20.21	抬頭	7 20.97	免用	1 10.42	盜竊	11 21.63	案件	22 17.59
取得	34 19.33	犯罪	21 20.69	全額	2 10.38	票載	3 20.76	憑證	8 17.17
募集	9 17.95	膨脹	7 19.8	小小	2 10.3	六獎	3 20.36	人因	8 17.02
印製	7 16.2	傳情	4 19.13	正規	2 10.2	開假	3 20.36	管理員	7 16.76
持	13 15.4	冒領	5 18.31	原	5 9.97	電腦版	3 19.02	收據	5 15.47
假造	3 14.93	盜領	5 18.16	欣榮	1 8.77	票券	11 17.95	魔方	2 15.45
發出去	2 14.13	捐給	5 17.81	作廢	1 7.66	開具假	2 17.84	普獎	2 14.52
買	9 13.66	開立	6 15.83	足額	1 7.0	預前	2 17.84	婚紗秀	2 14.07



SketchDiff: Comparing the behaviors of two words

WS Diff for 明星 and 演員 - Microsoft Internet Explorer

檔案(F) 編輯(E) 檢視(V) 我的最愛(A) 工具(T) 說明(H)

← 上一頁 → 搜尋 ☆ 我的最愛

網址(D) http://corpora.fi.muni.cz/chinese_all/run.cgi/wsdiff?corpname=chinese_all_trd;lemma=明星;pos=lemma2=演員

Google Bookmarks PageRank 49 blocked Check AutoLink AutoFill Send to Settings

明星/演員 chinese_all_trd freq = 23923/23213

[change options](#)

Common patterns

明星	21	14	7	0	-7	-14	-21	演員	
a_modifier	1729	4116	1.7	4.2	measure	688	1532	1.1	2.5
著名	64	736	39.6	70.4	位	310	576	50.3	54.0
武打	65	14	62.6	31.1	名	79	617	26.4	50.4
老牌	42	110	44.5	56.4	批	13	39	17.3	25.7
資深	6	237	11.2	53.5	個	79	159	21.3	23.9
知名	29	232	28.7	52.3	屆	13	7	12.6	5.7
曾經	43	36	49.4	38.7	場	11	5	12.5	4.9
大	468	19	46.6	5.1	次	9	17	8.1	9.5
年輕	8	136	13.3	43.8	possessor	459	921	1.0	2.1
最佳	27	135	23.2	39.5	知名度	6	10	17.6	21.1
小	38	232	20.5	37.9	國家	9	26	4.8	8.6
老	21	112	19.6	35.5	and/or	336	1517	0.4	1.7
眾多	44	16	33.8	17.5	導演	11	381	21.3	68.4
一流	9	22	18.5	25.4	歌星	16	50	32.2	42.6
名	11	198	4.6	25.3	歌手	7	50	17.0	34.8
新一代	12	13	22.9	20.6	藝術家	6	36	15.5	30.3
已故	11	7	22.5	14.9	n_modifier	13964	9897	1.6	1.2
					夫牌	162	13	57.7	21.6
					偶像	297	8	57.1	11.7
					喜劇	50	144	33.8	52.8
					影視	248	64	50.8	32.3
					好萊塢	190	70	50.5	37.0
					新生代	25	133	24.0	50.3
					男	41	355	16.1	46.6
					演技派	31	37	39.8	44.5
					歌仔戲	11	83	15.4	42.6
					芭蕾舞	7	63	13.6	42.5
					青年	26	564	5.4	41.7
					舞蹈	12	214	7.9	41.4
					電影	343	301	38.5	39.4
					歌劇	11	64	13.4	35.2
					男女	28	112	14.1	31.7
					女	234	18	30.9	7.2



Taiwan Presidential Corpus

- Containing speeches given by four Taiwan presidents on New Year's days and Double Tenth days since 1955 to 2007
- Over 230,000 characters in total
- Segmented and Tagged
- PI: Chu-Ren Huang and Kathleen Ahrens
- Data Collected by: Louis Lu

查詢結果

陳水扁	蔣經國	蔣介石	李登輝
207	20	188	43

顯示詞類 進階處理 重新查詢

共有 458 筆資料

510101 蔣介石 [more](#)在臺灣農村已經做到耕者有其田，普遍提高了**人民**的生活水準，[more](#)

510101 蔣介石 [more](#)只問是非，不計利害，為**人民**的生命自由，為國家的獨立完整與世界的和平[more](#)

510101 蔣介石 [more](#)蠻歷史，再加上俄帝對於實行恐怖統治，奴役**人民**的現代技術，其禍國殃民的罪惡，[more](#)

510101 蔣介石 [more](#)字改革」等等暴政，實際上就是要將我們大陸**人民**從精神至肉體，從個人至家庭，緊緊捆綁在俄[more](#)

510101 蔣介石 [more](#)中山先生的建國遺志，發揚我們先烈為國家與**人民**而獻身的救國精神，消滅共匪，打倒俄寇，[more](#)

510101 蔣介石 [more](#)全國軍民同胞：今年這一年將是全世界被奴役**人民**反共抗俄的革命高潮全面展開的一年。也就是[more](#)

510101 蔣介石 [more](#)且他們在蘇俄帝國主義操縱之下，完全將我們**人民**的生命和自由，[more](#)

510101 蔣介石 [more](#)底消滅。其一就是漢奸國賊惡貫滿盈終為全國**人民**所共棄，自必為民族革命勢力所摧毀；[more](#)

510101 蔣介石 [more](#)更指出了中國**人民**要救國自救，則唯一的主義是三民主義，唯一[more](#)

510101 蔣介石 [more](#)去寡廉鮮恥虛偽欺詐的陳腔爛調，來麻醉大陸**人民**，企圖緩和我大陸上的反共革命；他們並拿這[more](#)

510101 蔣介石 [more](#)消除「思想改造」的恐怖，以恢復**人民**思想、研究的自由。（四）取消「公私合營制[more](#)

510101 蔣介石 [more](#)」及對民生必需品的「統購」「統銷」等扼殺**人民**生計的苛政，維護私人合法財產及利潤，免除[more](#)

510101 蔣介石 [more](#)以恢復**人民**經濟生活的自由。（五）徹底消滅階級鬥爭，[more](#)

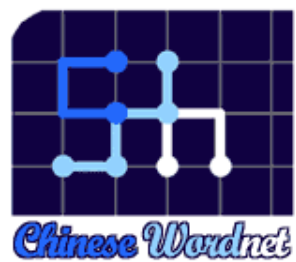
510101 蔣介石 [more](#)仇，互相殘殺，免除循環報復的恐怖，以恢復**人民**生命安全的自由。（六）發揚民族精神，[more](#)

510101 蔣介石 [more](#)道德，維護婚姻自由，增進家庭康樂，並保障**人民**言論、出版、集會、結社、居住與宗教信仰之[more](#)

510101 蔣介石 [more](#)免除一切控制、迫害的恐怖，以恢復**人民**選擇生活方式的自由。至於大陸的匪黨分子、[more](#)

70101 蔣介石 [more](#)經事實證明：一年來共產鐵幕內八億**人民**自東歐到東亞，到處爆發了反共抗暴的運動；[more](#)

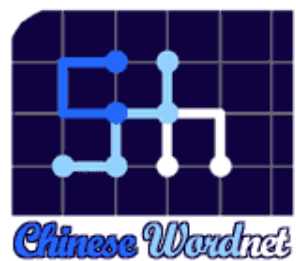
70101 蔣介石 [more](#)命克星，就是鐵幕內與其不共戴天的八億反共**人民**，[more](#)



Language Resources as Research Infrastructure

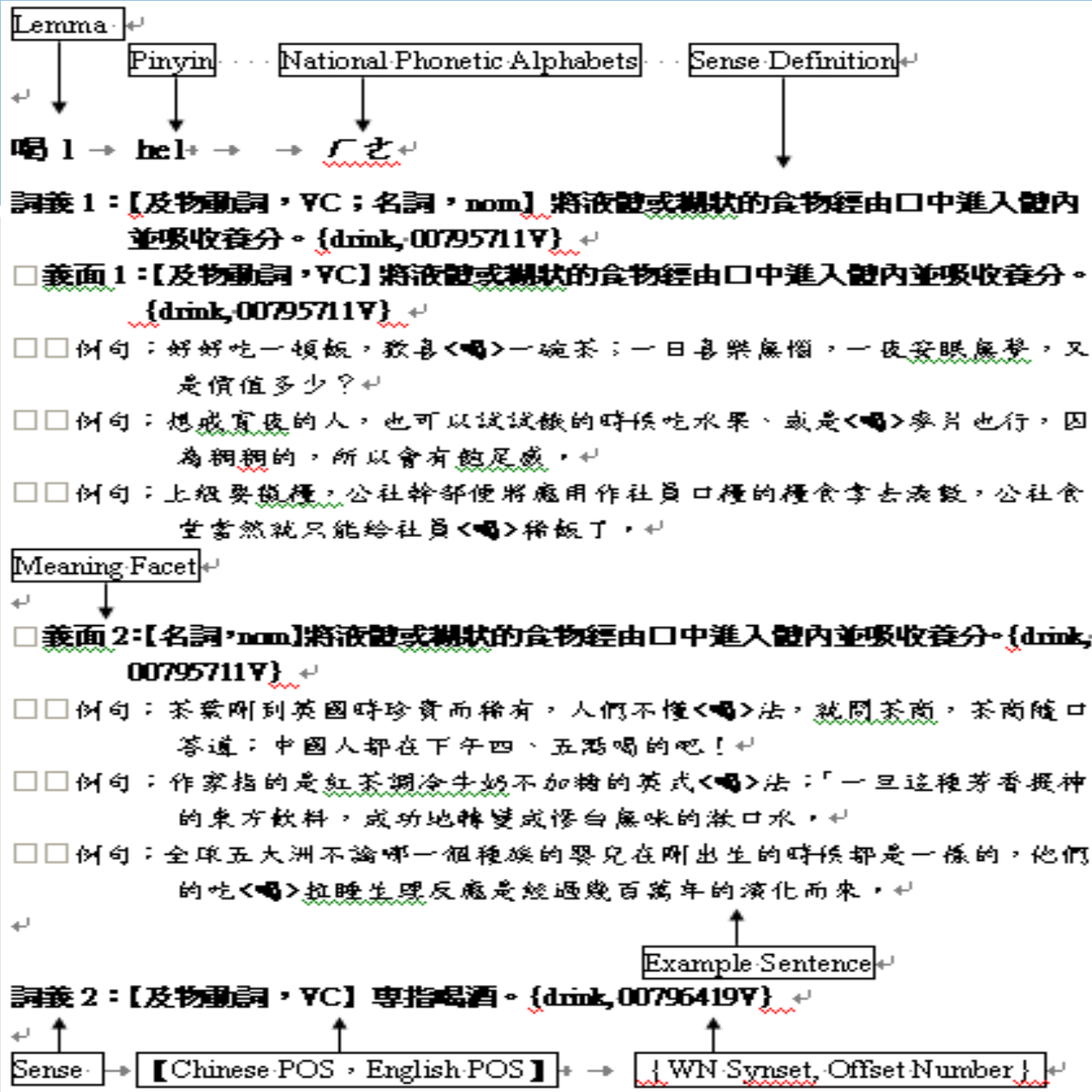
- WordNet/Ontology
 - Chinese Wordnet
 - Sinica BOW
 - Hantology

語音



CWN: Chinese Wordnet

- 8,000 lemmas,
- 20,000 senses and 5,000 facets
- Each sense is attested by at least 4 corpus sentences
- Each attested sense has up to 20 sentences manually sense-tagged in Sinica Corpus
- Semi-automatic sense-tagging of complete corpus in process



<http://cwn.ling.sinica.edu.tw/>

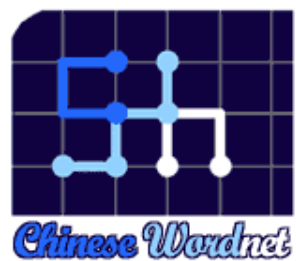
The screenshot shows a Microsoft Internet Explorer browser window displaying the Chinese Wordnet website. The browser's address bar shows the URL <http://cwn.ling.sinica.edu.tw/>. The website header includes the text "Chinese Wordnet" and "中文詞彙網路". A navigation menu on the left lists several links: 簡介, 詞義描述規範, 領域標記表, 辭彙網路查詢, 網路資源, and 研究小組. The main content area features two buttons: "詞彙查詢" and "詞類索引". Below these buttons, there is a text input field for keywords and a radio button menu for search scope. The radio buttons are: 中文詞彙 (selected), 釋義內文, 英文對譯, 中文詞彙模糊查詢, 注音, and 漢語拼音. At the bottom of the search area, there are "送出" and "重填" buttons.

June



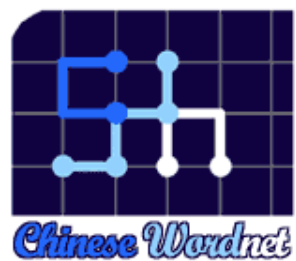
Lexical Knowledgebase

Core Foundation
for Language Acquisition

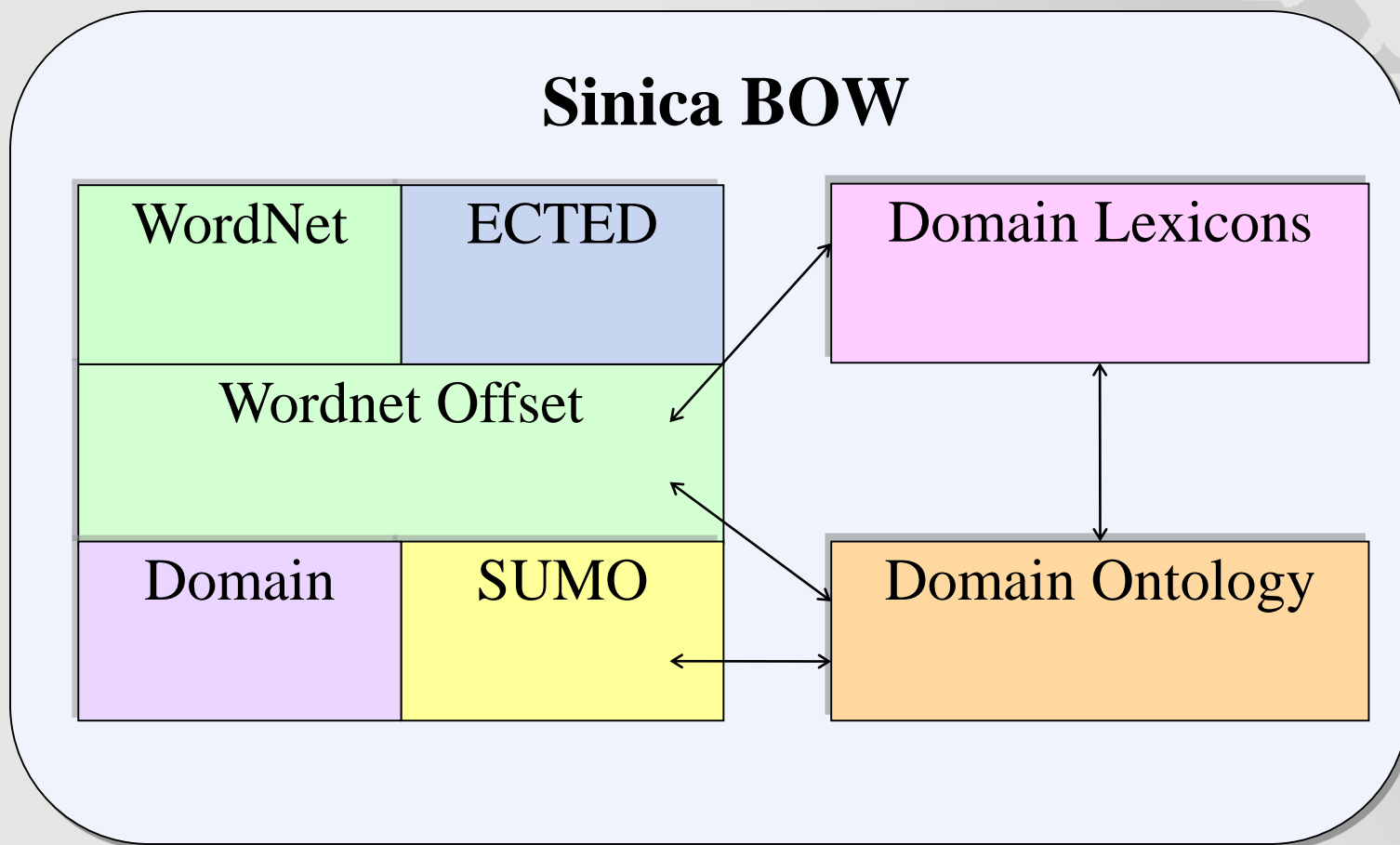


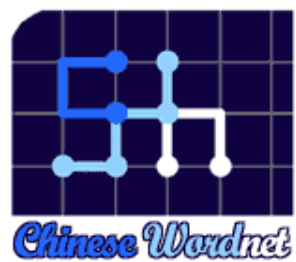
Sinica BOW

- **Academia Sinica Bilingual Ontological Wordnet**
- 2003
- Website:
 - <http://bow.sinica.edu.tw/>



The resource and structure of Sinica BOW





Linking linguistic Information to SUMO Ontology

Machine - Microsoft Internet Explorer

網址: http://virtual.cvt.cz/kfb-test/ont/concepts/_machine.html

機器 (Machine)

Machines are Devices that that have a well-defined resource and result and that automatically convert the resource into the result.

Ontology

SUMO / OBJECTS

Superclass(es)

- ↑ 實體
- ↑ 物質的
- ↑ 物體
- ↑ 自身進類物體
- ↑ 複雜體/非單質體
- ↑ 人造物
- ↑ 裝置
- ↑ 機器

Coordinate term(s)

工程元件 樂器 運輸工具 武器

Related WordNet synsets

中央研究院中英雙語知識本體詞網 The Academia Sinica Bilingual Ontological Wordnet (BOW)-知識本體 Ontology - Microsoft Internet Explorer

網址: http://bow.sinica.edu.tw/ont/

知識本體 Ontology

最新消息 系統簡介 詞網 知識本體 索引 檔案櫃 相關連結 回首頁

建議上層共用知識本體 SUMO 領域知識本體 Domain Ontology

以詞彙查詢SUMO LookUp Word in SUMO

Keyword 檢索條件值: computer

Sense1: an expert at calculation (or at operating calculating machines)

SUMO Equivalents	SUMO English Concept	SUMO 中文概念
Human	human	人類

Sense2: a machine for performing calculations automatically

SUMO Equivalents	SUMO English Concept	SUMO 中文概念
Machine	machine	機器

共查到2個對譯詞義 Word LookUp 以詞彙查詢SUMO

- : hypernym, 表示SUMO的 Concept包含其同義詞集
- : synonymy, 表示SUMO的 Concept和其同義詞集相等
- : instantiation, 表示SUMO的 Concept既不相等, 也不包含其同義詞集

June 30, 2008

Inst. of Ling., Academia Sinica

Query by Domain

中央研究院中英雙語知識本體詞網 The Academic Store Bilingual Ontological WordNet (BOW)-索引 Index - Microsoft Internet Explorer

http://bow.sinica.edu.tw/ind/

索引 Index

最新消息 系統簡介 詞網 知識本體 索引 臺灣廳 相關連結 回首頁

字詞索引 字彙索引 字尾索引 詞類索引 來源索引 頻率索引 領域索引 概念索引 特色索引

Nca (地方詞), WordNet中英對譯, Core Word (核心詞彙), 中文綜合索引 第1頁 查詢LookUp

九州	三星	上議院	土耳其	大不列顛
大內	大安	大西洋	大阪	大洋洲
山東	中美	中國	中華民國	公安局
匹茲堡	天山	天津	天津	太平洋
太陽系	巴西	巴拉圭	巴林	巴拿馬
巴勒斯坦	巴基斯坦	巴黎	日內瓦	日本
比利時	火星	牙買加	以色列	加拿大
加納比	加納比海	北大	北京	北齊
北海	北海道	北國	北極圈	北歐

Chinese/English Search

WordNet E-C corres.

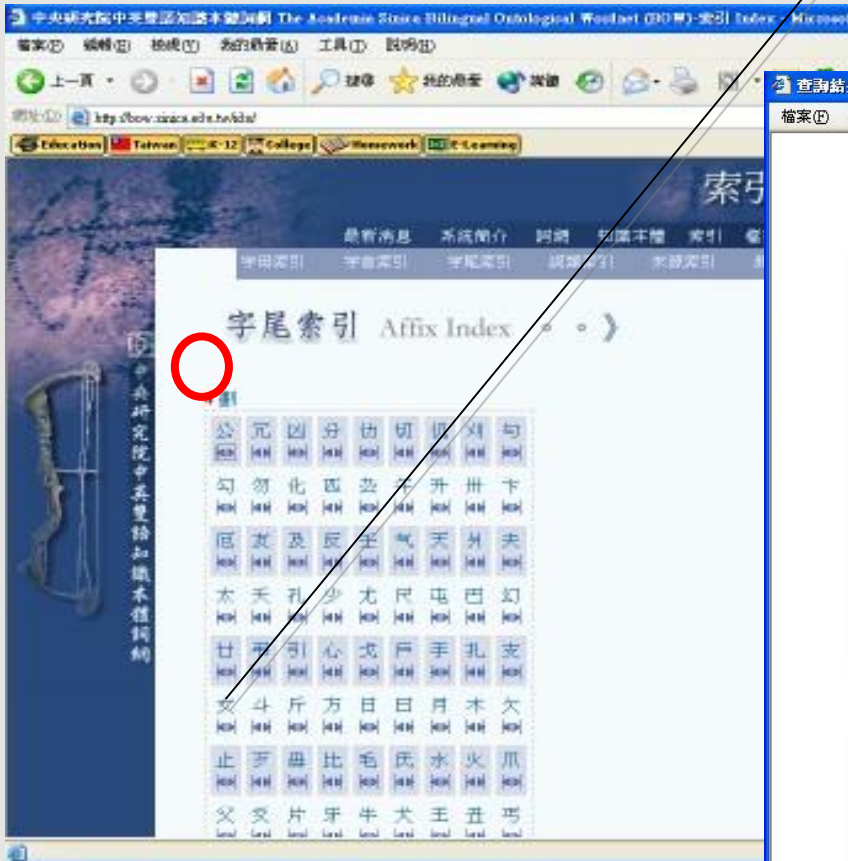
SUMO Lookup

June 30, 2008

Inst. of Ling., Academia Sinica

Query by Prefix or Suffix

List of Chinese prefixes and suffixes



字尾索引 Affix Index

公	元	因	分	切	刈	勾
勾	勿	化	西	志	升	卅
卮	友	及	反	壬	气	天
夫	太	天	孔	少	尤	尺
屯	巴	幻	甘	引	心	戎
尸	手	扎	支	文	斗	斤
方	日	日	月	木	欠	止
歹	母	比	毛	氏	水	火
爪	父	爻	片	牙	牛	犬
王	丑	丐				



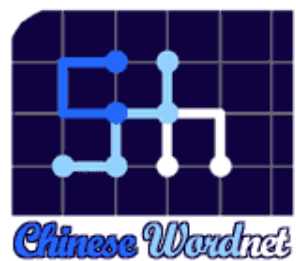
查詢結果

你查詢的 **公** 在詞首字出現的情形如下：

詞義 1	詞類	A	詞林類	Ed551	詞林領頭詞	公
	詞義	public, state-owned				
	構詞律	例子個數	26			
		例子				
		公安	公式	公序	公決	公私
		公車	公物	公信	公保	公勞
		公庫	公時	公證	公自	公費
		公義	公路	公僕	公倍	公數
		公誼	公膳	公錢	公職	公關
		公權	公廨			

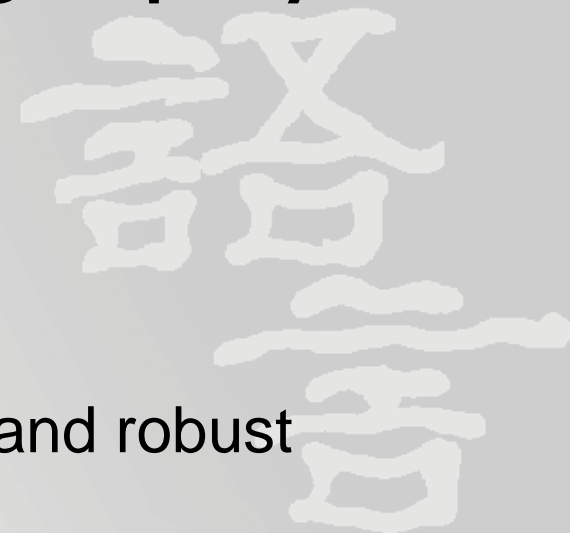
你查詢的 **公** 在詞尾字出現的情形如下：

詞義 1	詞類	Nab	詞林類	Ab011	詞林領頭詞	男人
	詞義	respectful form of address for elderly man				
	構詞律	例子個數	9			
		例子				
		姊公	家公	廟公	帳公	大豬
		老祖	孝廉	四叔	表叔	

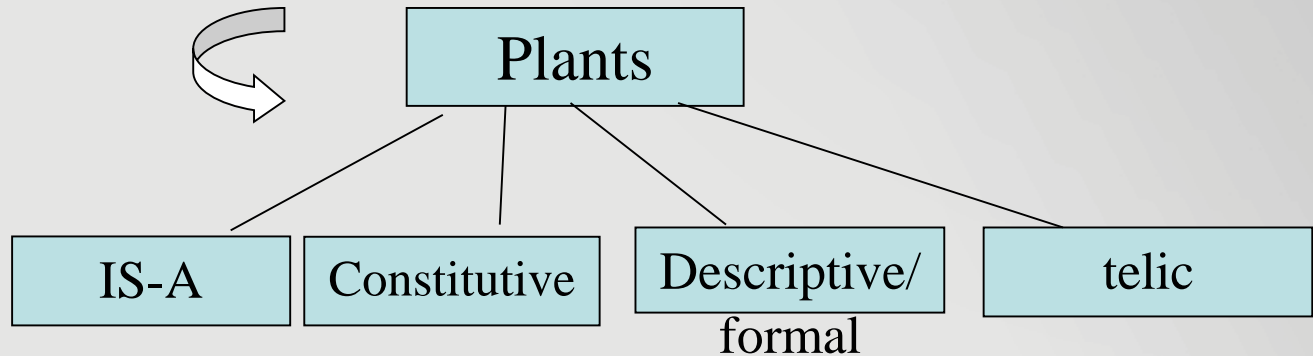
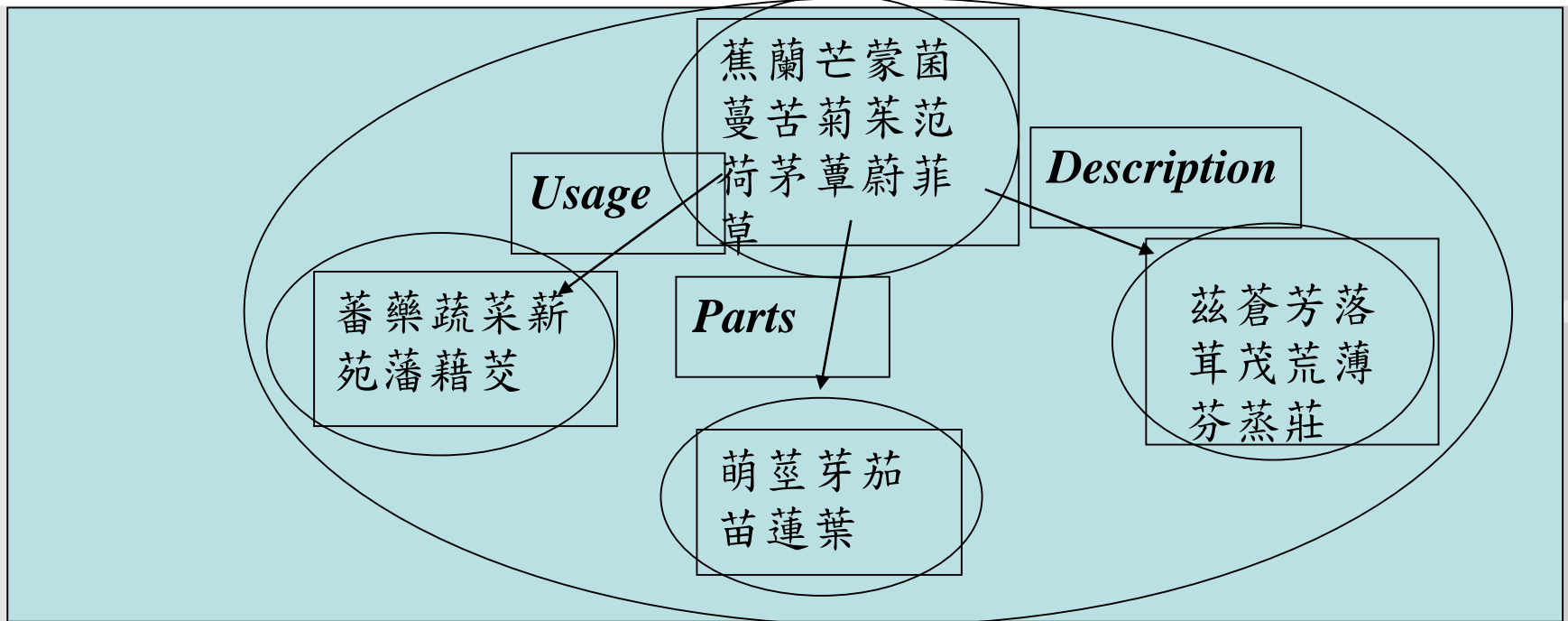


Ontologies Based on Chinese Orthography

- Hantology (Ya-Min Chou 2005)
- HanziNet (Shu-Kai Hsieh 2005)
- Our Common Goal
 - To utilize the highly conventionalized and robust Chinese Orthography
 - As well as its concept-based classificatory System explicitly represented by radicals
 - To extract conceptual system and build linguistic ontology for knowledge processing and conceptual explanation

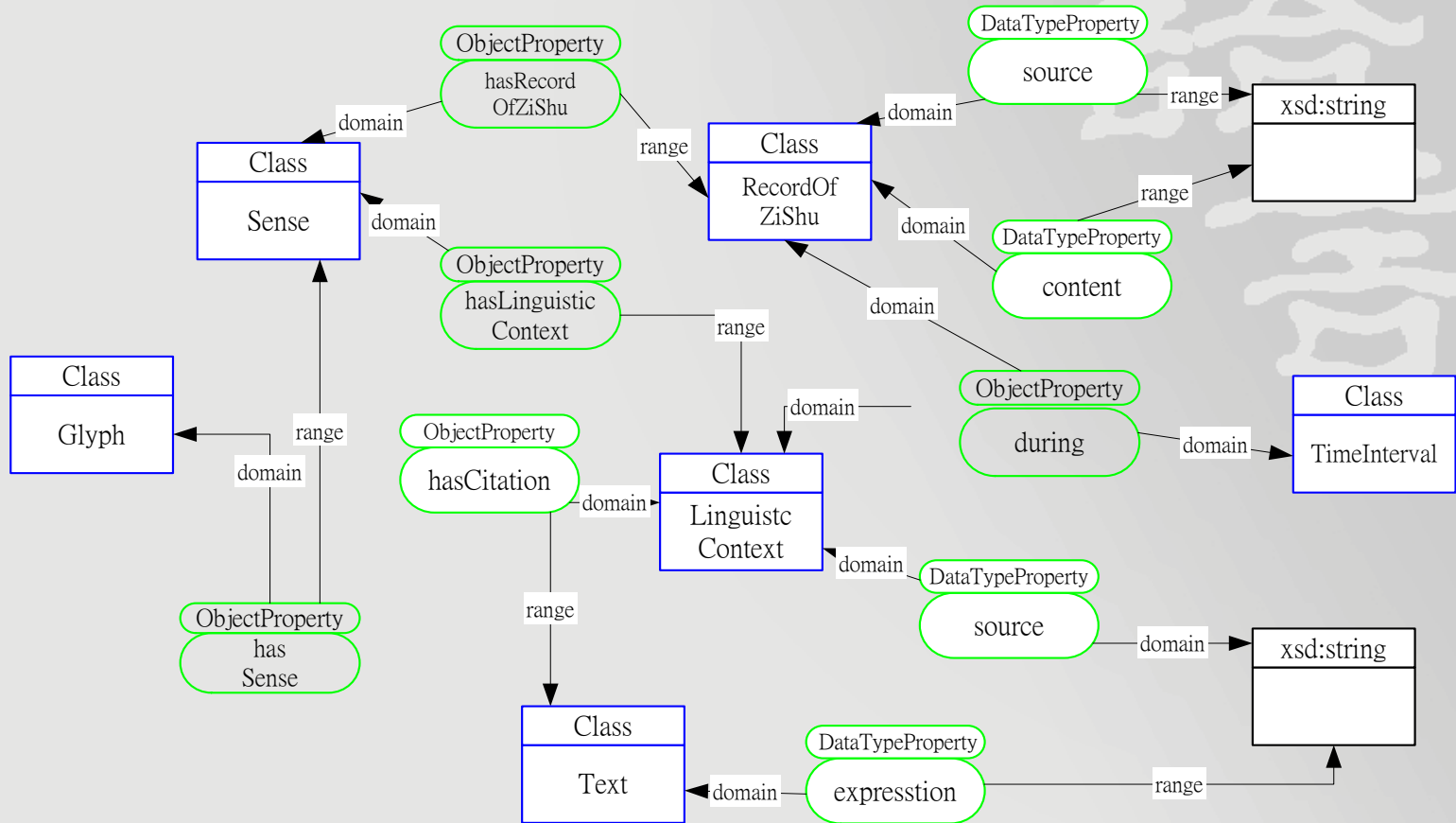


Hantology: Character and Knowledge System





MANtology for the Semantic Web: OWL Representation



June 30, 2008

Inst. of Ling., Academia Sinica

Hantology Browser (Ya-Min Chou)

Hantology 1.5
 知識本體(O) 檢視(V) 視窗(W) 檢索 說明(H)

字形關係

- 臭
 - 解形(Composition)
 - 形符聲符
 - 六書(Principle of formation): 會意
 - 構字式(glyph expression): 自+犬
 - Glyph Evolution
 - 小篆
 - 臭
 - Sense: 用鼻子聞氣味[本義]
 - POS: V動詞
 - Pronunciation
 - Ancient Pronunciation
 - Middle Pronunciation
 - Modern Pronunciation
 - Taiwan
 - 注音一式: ㄉㄡˋ ㄨㄛˋ 4
 - Citation and Context
 - 解釋形義
 - 例證
 - 戰國: 荀子
 - 成事之俎不嘗也, 三臭之不食
 - Variant 異體
 - 噏
 - 鼻
 - 鼻
 - 嗅
 - Time 時間
 - 初見於: 戰國
 - Sense: 難聞的氣味[弓|申義]
 - Pronunciation
 - Ancient Pronunciation

字義概念

- 臭
 - 主要意符(main semantic-symbol)
 - 犬
 - 初義(original sense)
 - 用鼻子聞氣味
 - 說文: 禽走臭而知其跡者, 犬也
 - 常用義(modern sense)
 - 現代常用義(modern sense) 1
 - 難聞的氣味
 - SUMO concept: 嗅覺屬性
 - 衍生詞彙(generated words): 三國: 臭、唐: 臭氣、
 - 現代常用義(modern sense) 2
 - 惡名, 不好的名聲
 - SUMO concept: 主觀評價屬性
 - 衍生詞彙(generated words): 元: 遺臭萬年
 - 現代常用義(modern sense) 3
 - 令人厭惡的
 - SUMO concept: 主觀評價屬性
 - 衍生詞彙(generated words): 明: 臭、現代: 臭事、
 - 現代常用義(modern sense) 4
 - 狠狠的
 - SUMO concept: 主觀評價屬性
 - 衍生詞彙(generated words): 清: 臭罵
 - 現代常用義(modern sense) 5
 - 氣味
 - SUMO concept: 嗅覺屬性
 - 衍生詞彙(generated words): 春秋: 臭、現代: 無臭
 - 衍生詞彙(generated words)
 - 出現在詞首(prefix)
 - 臭小子、臭水溝、臭皮囊、臭名、臭吃臭喝、臭
 - 出現在詞尾(suffix)

意符概念

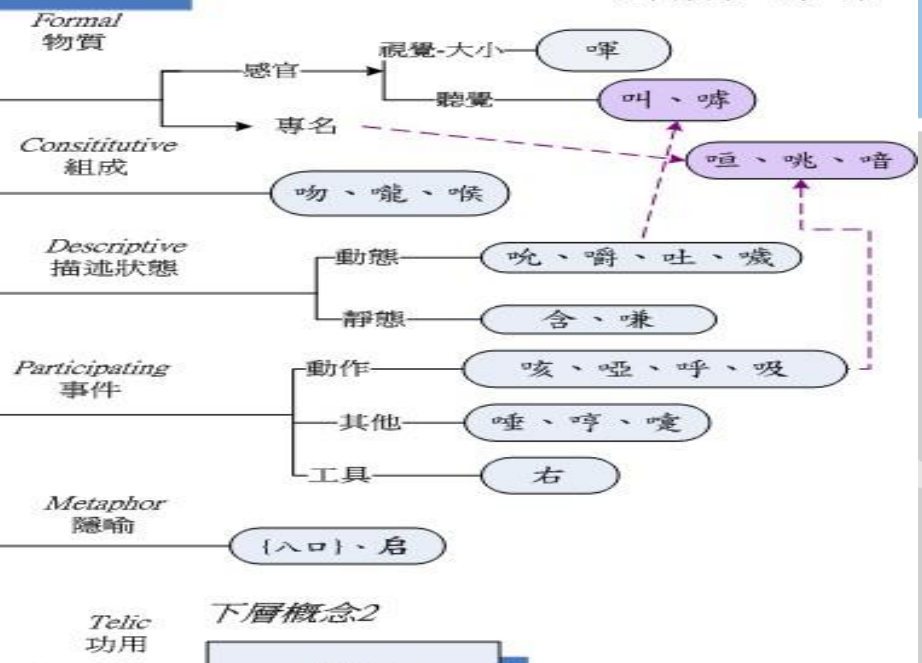
- 意符(semantic-symbol)
 - 犬
 - 概念(concept): 狗
 - SUMO concept: 犬科動物
 - 自
 - 概念(concept): 鼻子
 - SUMO concept: 器官

狀態: 就緒

口

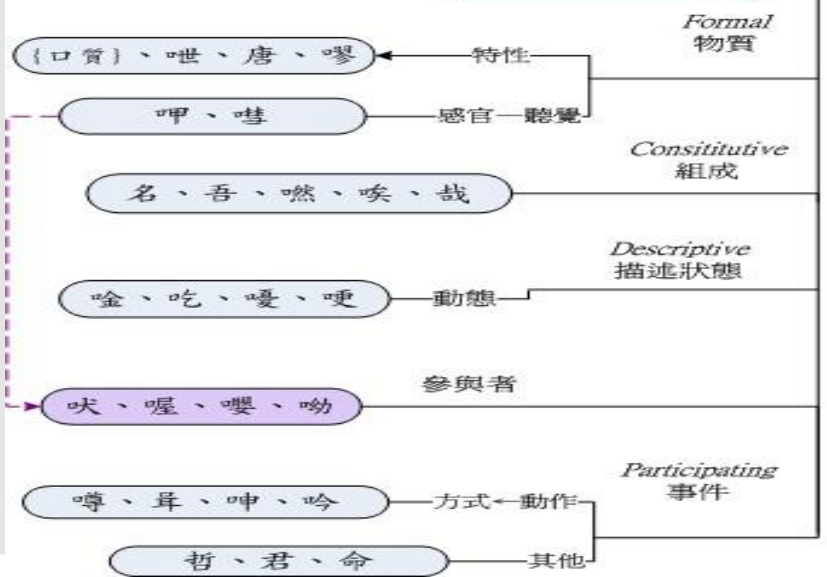
基本概念——用於言語和飲食的身體器官

※同義字：囁、喙



下層概念1

言語



下層概念2

飲食

