




Guangwai Lancaster Chinese Learner Corpus

CHINESE APPLIED CORPUS LINGUISTICS SYMPOSIUM

Lancaster University, 29 April

Answers to both simple and complex questions



Simple query: Make Concordance

Query Tools Context Text Types

Text types

Subcorpus: create new

TEXT.MODE: speech writing

TEXT.GENRE: 口头采访 口语考试 口语课堂练习 口语辅导 会话范文 会话范文

TEXT.INTERACTION: HA dialogue monologue

TEXT.TOPIC:

U.ROLE: examiner native_speaker nonnative_speaker student

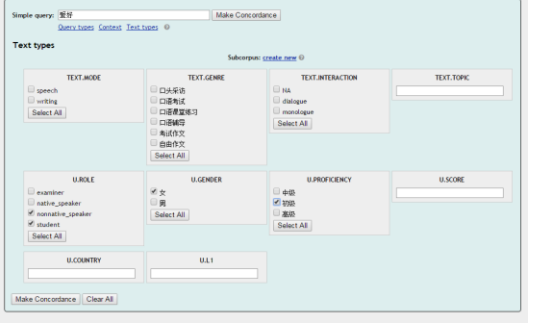
U.GENDER: 女 男

U.PROFICIENCY: 中低 中 中高

U.SCORE:

U.COUNTRY:

U.L1:



Simple query: 提问 Make Concordance

Query Tools Context Text Types

Text types

Subcorpus: create new

TEXT.MODE: speech writing

TEXT.GENRE: 口头采访 口语考试 口语课堂练习 口语辅导 会话范文 会话范文

TEXT.INTERACTION: HA dialogue monologue

TEXT.TOPIC:

U.ROLE: examiner native_speaker nonnative_speaker student


U.GENDER: 女 男

U.PROFICIENCY: 中低 中 中高

U.SCORE:

U.COUNTRY:

U.L1:



Home

Concordance

Query: 提问, 女, 初级, nonnative_speaker, student, 51 (sorting: ...)

Page 1 of 1 [Go] Next Last

Word sketch

Thesaurus

Search all

Context info

Manage corpus

My jobs

User guide/CF

Save

Make subcorpus

View options

Kwic

Sentence

Sort

Left

Right

Node

References

Shuffle

Filter

Rich meta-data

Rich meta-data

Query 爱好, 女, 初级, nonnative...

Page 1 of 6 Go Next

student.多哥... 爸爸是医生, 我

student.多哥... 医生, 呢, 我

student.多哥... 爸爸爱好走路

student.俄罗斯... 吗? 我觉得学

student.高尔... 水果? 我喜欢吃

student.安哥拉.初级.葡语.女

student.哈萨... 有什么爱好

student.哈萨... 你, 你, 有

student.哈萨... 你有什么爱好

student.哈萨... 十年, 呢, 我

student.哈萨... 呢, 听说

student.哈萨... 有什么爱好

student.西班牙... 呢, 我想

student.意大利... Yes 呢

student.意大利... 的爱好”

student.意大利... 。呢。呢

student.意大利... 去在桂林

student.意大利... 我去在意大利

student.意大利... 呢, 我爱好

student.意大利... 买衣服。我

Page 1 of 6 Go Next

Comparison with native speaker corpora

Word list

Corpus: Guangxi_Lancaster_2.7

Subcorpus: nns

Reference corpus: LCMC

Switch between reference and subcorpus

Page 1 of 1 Go Next

word	Guangxi_Lancaster_2.7: nns		LCMC	
	Freq	Freq/mill @	Freq	Freq/mill
呢	20,612	19921.1	6	5.9
爱好	1,238	896.1	0	0.0
中国人	1,087	847.3	0	0.0
呢	2,283	7255.3	10	9.8
住在	669	516.7	0	0.0
广外	560	432.4	0	0.0
朋友们	538	412.5	0	0.0
男的	363	280.4	0	0.0
不大	352	275.8	0	0.0
OK	352	275.8	0	0.0
南方	321	247.9	0	0.0
爸爸	226	173.2	0	0.0
到了	220	169.4	0	0.0
不知道	220	169.4	0	0.0
还学	245	189.2	0	0.0
那时候	220	181.5	0	0.0
去过	208	160.7	0	0.0
同学们	204	157.4	0	0.0
韩国人	203	156.8	0	0.0
字了	198	152.8	0	0.0
在一起	187	145.2	0	0.0
不是	388	299.7	1	1.0
聊天儿	188	145.2	0	0.0
不一样	186	143.7	0	0.0
如果	184	142.1	0	0.0
那她	181	139.8	0	0.0
旅行	2,132	1650.7	11	10.8
美国人	179	137.5	0	0.0
阿尔	171	131.1	0	0.0
看书	120	91.3	0	0.0
冲杯茶	144	111.4	0	0.0

Comparison with The Lancaster Corpus of Mandarin Chinese

Other things...

General discussion

1. Do you have experience with learning/teaching Chinese? What are the main challenges?
2. Can corpus linguistics address some of these challenges? How?
3. What areas of research should we focus on as a priority? Are there any 'burning questions' in the field?