



A U.S. Department of Education Title VI National Language Resource Center at the Pennsylvania State University © CALPER 2012. All Rights Reserved.

L2 learner corpus data by courtesy of Dr. Jie Zhang, University of Oklahoma

CALPERLEX

Chinese

Learner Corpora

Contributed by Haiyang Ai (Penn State), supervised by Dr. Xiaofei Lu (Penn State).

Analyzing learner corpus data with GOLD: The use of the aspectual marker *le* (了) in L2 Chinese writing

Language teachers can create, query, and analyze corpora of language samples produced by their students using the tools available in GOLD*. In this document, we illustrate how GOLD can be used to examine L2 Chinese learners' use of the perfective aspectual marker *le* (了) in writing.

I. Overall frequency

Querying for the overall frequency of a particular linguistic structure in a corpus is straightforward in GOLD. After selecting the corpus to search, one simply enters the keyword in question. As is shown in the figure below, our search for the perfective marker *le* (了) returns 95 instances in 32 documents.

Results:		
Showing 95 results of 95 in 32 documents		
#	Concordance:	View:
1	不亦乐乎！”这儿家乡费城我最喜欢费城了。费城有很多人，很大，还有便	View
2	我只是吃了几口，我便吃了。今天真高兴，“有朋自远方来，不亦乐乎！	View
3	我们都吃了很多。虽然我只是吃了几口，我便吃了。今天真	View
4	饭很好吃，色香味俱全。我们都吃了很多。虽然我只是吃了几口，	View
5	在那儿。这儿比我白家乡有很多了农田了。这儿像那儿发展得一样决	View
6	。这儿比我白家乡有很多了农田了	View

II. Frequency of use by a focal learner

It is possible to zero in on a particular student. In order to do that, we need to expand the query to include the relevant meta-information. For example, when we specify the search condition that *Student Number* equals *S1*, the search for *le* (了) returns 4 instances in 2 documents produced by *S1*.

Center for Advanced Language Proficiency Education and Research

Results:
Showing 4 results of 4 in 2 documents

#	Concordance:	View:
1	的生活。我在宿舍主了三个星期了。我们正忙者做中文的功课。	View
2	很喜欢大学的生活。我在宿舍主了三个星期了。我们正忙者做中文	View
3	几个航空公司的网站找得到便宜的机票了。如果你要去，我可以马上订机票	View
4	，跟我一起去北京旅游。我查了几个航空公司的网站找得到便宜的机票	View

III. Longitudinal development

If the corpus contains samples produced by students at different time points, it is possible to examine the learners' longitudinal development. For example, the L2 Chinese Written Corpus contains essay written at three time points, i.e., 09/2009, 10/2009, and 11/2009. Now to examine longitudinal development of *S4*, we can specify 1) the keyword – *le* (了), and 2) the focal student – *S4*, and 3) a temporal reference – one of the three time points, for each query and then repeat the search with each of the other time points.

09/2009

Results:
Showing 6 results of 6 in 2 documents

#	Concordance:	View:
1	。在那个考试的时候，我不少紧张了。我写字写得很慢。不过，	View
2	一个交叉口。在那个交叉口往左拐。过了三个红绿灯，就友边你会看到我的家	View
3	它。昨天我的同学们和我有了一个试考得很难。在那个考试	View
4	，请放心。在中文课我把累死了。但是我爱的它。昨天我的同学	View
5	？我已经上大学住了三个月多了。一切都很好，请放心。在	View
6	？最近好吗？我已经上大学了三个月多了。一切都很好，	View

10/2009

Results:
Showing 10 results of 10 in 1 documents

#	Concordance:	View:
1	不亦乐乎！”这儿家乡费城我最喜欢费城了。费城有很多人，很大，还有便	View
2	在那儿。这儿比我白家乡有很多了农田了。这儿像那儿发展得一样决	View
3	。这儿比我白家乡有很多了农田了。这儿像那儿发展得一样决。那儿	View
4	。那儿没有这儿那么旅客。这有多了旅馆。	View
5	我只是吃了几口，我便吃了。今天真高兴，“有朋自远方来，不亦乐乎！	View
6	我们都吃了很多。虽然我只是吃了几口，我便吃了。今天真	View
7	。这儿的天气一天比一天冷了。不能吃我妈妈的家常便饭我不习惯	View
8	习惯。上个星期三我的姐姐来拜访了我。每天都我跟她一起出去吃饭	View
9	饭很好吃，色香味俱全。我们都吃了很多。虽然我只是吃了几口，	View
10	。我每次回家的时候，妈妈特地做了一些拿手菜。这儿的天气一天比一	View

11/2009

Results:

Showing 7 results of 7 in 2 documents

#	Concordance:	View:
1	韩国坐船走得到中国。我以前听过了。你会用United Airlines航空公司去北京。	View
2	。你想看风景的时候，火车好多了。在北京我们应该带地图。有的地图	View
3	上查一查了。我在中国Yahoo上查了。以前我听说北京有多美丽的地方	View
4	想去中国的北京。我网站上查一查了。我在中国Yahoo上查了。以前	View
5	。最新的特技效果常常比以前的好多了。一编来一编人物的描写也	View
6	编来一编人物的描写也进步了。	View
7	从前我跟《三国志》玩过。KOEI是 日本公司。KOEI是 做 这个作品	View

IV. Qualitative analysis of erroneous use of *le*

For each file in the learner corpus, GOLD allows the user to examine, or “view” the content of the file, as well as any meta-data encoded.

Close

Metadata

Character: Simplified	Essay number: 150
Semester: 3	Student number: S9
Time: 10/2009	title: 003150E6_S_S9
Topic: A thank you email	

Contents

谢谢你们给我一个生日宴会。它把我很高兴，你们的礼物都很体贴。我的蛋糕色香味俱全。我不要吃永远。谁雇佣表演者。我喜欢了。我希望你们都可以再回明年。我知道我的家不怎么大。我怪你们都适合了。老人谢谢为来了。我知道你们的家挺远。我觉得我是一年比一年老了。可是你们促使我觉得年轻。昨天晚上的时候，你们者给我干杯挺好。我浩叹被你们的仁慈。现在我有很多东西。我要把家全正理完。在你们的生日的时候我要给你们礼物好的。我太高兴所以我睡不着。明天我会很困。再见。

To identify erroneous uses, we first searched for *le* (了), and then manually examined the concordance lines to identify incorrect uses. Two examples of erroneous uses are given below.

- (1) 谁雇佣表演者。我喜欢了。

shui guyong biaoyanze, wo xihuan le

who hire performer, 1SG like PRF

“Who hired the performer? I like (it).”

(S9, 10/2010)

Center for Advanced Language Proficiency Education and Research

(2) 因为以前我想了成律师。

yinwei yiqian wo xiang le cheng lvshi

because before 1SG want PRF become lawyer

“Because I wanted to be a lawyer before”

(S3, 11/2009)

In both (1) and (2), *le* is unnecessary. In (1), the student was expressing her appreciation for the birthday party her friends organized for her. It is grammatically incorrect to use the perfective marker *le* immediately after the verb 喜欢 “like”, which denotes a mental state.

In (2), the student was describing his career plan in the past. In Chinese, the notion of past tense is realized through lexical words, in this case, *yiqian* “before”. The perfective *le* is unnecessary because this sentence does not indicate that his past career plan has been fulfilled.

*The Graphic Online Language Diagnostic (GOLD) is a web-based easy-to use assessment instrument. It enables language teachers to create their own corpora from their students' spoken and written performance and to analyze and benchmark student performance. It also allows teachers to trace the development of individual students or groups of students over time.

GOLD was developed by CALPER faculty and staff. It is funded by a grant from the U.S. Department of Education (Title VI, P229A060003 and P229A020010) to the Center for Advanced Language Proficiency Education and Research and through support from the Center for Language Acquisition at Penn State. Its use is free to language educators, but registration is required. URL: <http://gold.gwserver1.net>