

# 外语教师的 数据素养与专业发展

中山大学 金 檀

上海交通大学 郭 凯 有米科技 李百川

# 分享提纲

- 数据素养**要素**
- 专业发展**案例**
- 专业发展**建议**

(一)

外语教师数据素养的  
构成要素



**THE BIG IDEA:** When students, parents, educators, and partners have the right information to make decisions, students excel.

**PRODUCED BY**  
**QDC**  
**DATA QUALITY**  
**CAMPAIGN**  
dataqualitycampaign.org

**What does it mean to use data in service of student learning?**  
 Data is one of the most powerful tools to inform good decisions and create opportunities for students along their education journey—and it's much more than test scores. Data helps us make connections that lead to insights and improvements. Everyone has an important role to play in helping all students succeed in their own individualized ways. Here's what it will look like when data is working for all students.

**STUDENTS**  
 "I know my strengths and where I need to grow. I can shape my own education journey."

**TEACHERS**  
 "I know where my students are succeeding and struggling right now."  
 "I can help them grow."

**SCHOOL LEADERS**  
 "I know what's working and what isn't in my school."  
 "I can make timely decisions and make sure resources support great teaching and improve student learning."

**PARENTS**  
 "I know what actions to take to help my child on her path to success."  
 "I can be a better champion for her."

**AFTERSCHOOL PARTNERS**  
 "I know what's happening with these kids before 3:00 p.m."  
 "I can help families and communities create more opportunities for students to succeed."

Many educators and policymakers are already taking steps in the right direction. See QDC's Four Policy Priorities to Make Data Work for Students for more on making this vision a reality for all students.



*“Because if science is a search in the darkness, then the humanities are a candle that shows us where we’ve been and the danger that lies ahead.”*





ELSEVIER

Contents lists available at ScienceDirect

# Teaching and Teacher Education

journal homepage: [www.elsevier.com/locate/tate](http://www.elsevier.com/locate/tate)



## What does it mean for teachers to be data literate: Laying out the skills, knowledge, and dispositions

Ellen B. Mandinach <sup>a,\*</sup>, Edith S. Gummer <sup>b</sup>

<sup>a</sup> WestEd, United States

<sup>b</sup> Ewing Marion Kauffman Foundation, United States



### ARTICLE INFO

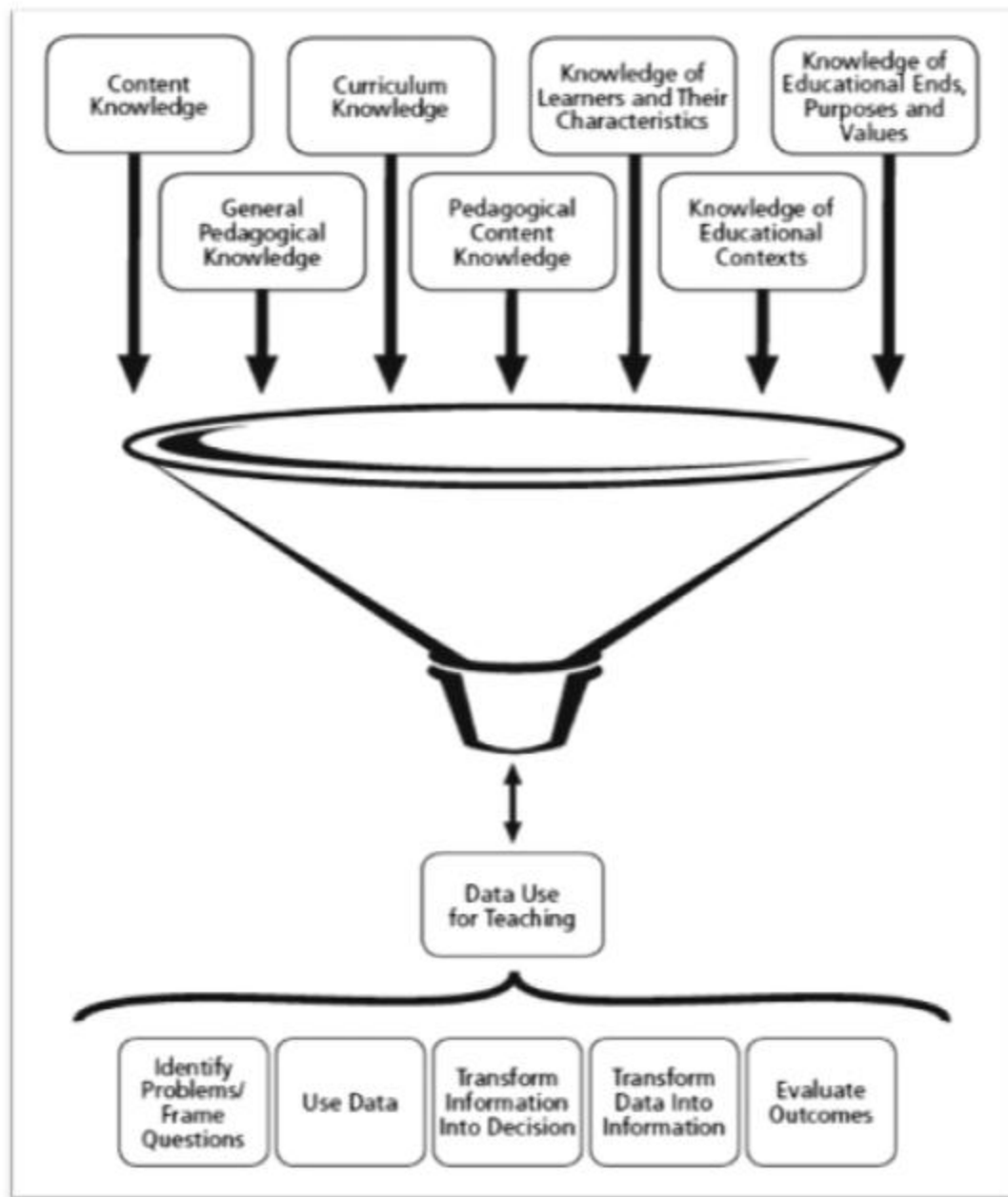
*Article history:*  
Received 3 February 2016  
Received in revised form  
7 July 2016  
Accepted 8 July 2016  
Available online 11 August 2016

*Keywords:*  
Data literacy  
Data-driven decision making  
Teacher knowledge

### ABSTRACT

This article reports on the evolution of a conceptual framework for a construct called data literacy for teachers. Data use has become an emphasis in education but few educators have received sufficient training or preparation pertaining to data literacy skills. This article lays out the framework, identifying the specific knowledge, skills, and dispositions teachers need to use data effectively and responsibly. It concludes with a call to schools of education and teacher preparation programs to begin to integrate data literacy into curricula and practical experiences.

© 2016 Elsevier Ltd. All rights reserved.



Xiaofei Lu

# Computational Methods for Corpus Annotation and Analysis

 Springer

tesol QUARTERLY

tesol.org

## INVITED TEACHING ISSUES

TESOL Quarterly publishes brief commentaries on aspects of English language teaching.

Edited by PHAN LE HA

*University of Hawaii at Manoa*

### *A Data-Driven Approach to Text Adaptation in Teaching Material Preparation: Design, Implementation, and Teacher Professional Development*

TAN JIN

*Sun Yat-sen University  
Guangzhou, China*

XIAOFEI LU

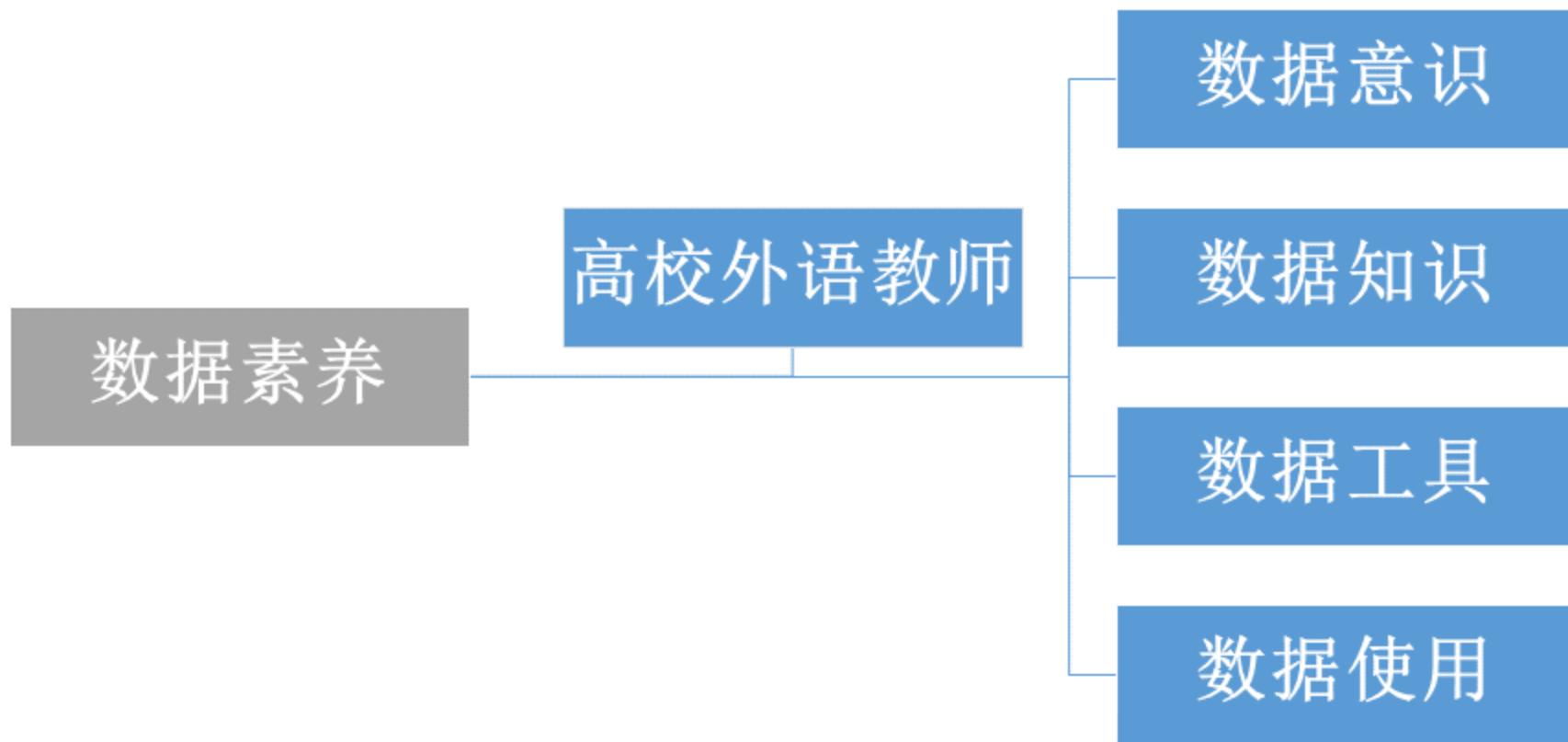
*The Pennsylvania State University  
University Park, Pennsylvania, United States*

doi: 10.1002/tesq.434

TESOL teachers play an indispensable role in teaching material preparation (Bunch, Walqui, & Pearson, 2014). A critical consideration in this process that directly bears upon learners' literacy development is to ensure the linguistic complexity of the reading texts is suitable for their expected proficiency levels (Crossley, Allen, & McNamara, 2012; Hiebert & Mesmer, 2013). Matching texts to particular proficiency levels is challenging even for experienced teachers (Sung, Lin, Dyson, Chang, & Chen, 2015). With the advent of corpus techniques, numerous data-driven indices are now available for characterizing text complexity (Crossley, Greenfield, & McNamara, 2008; Graesser, McNamara, & Kulikowich, 2011; Lu, 2014). While not every aspect of text complexity (e.g., topic familiarity and cultural

TESOL QUARTERLY Vol. 0, No. 0, xxxx 2017  
© 2017 TESOL International Association

1

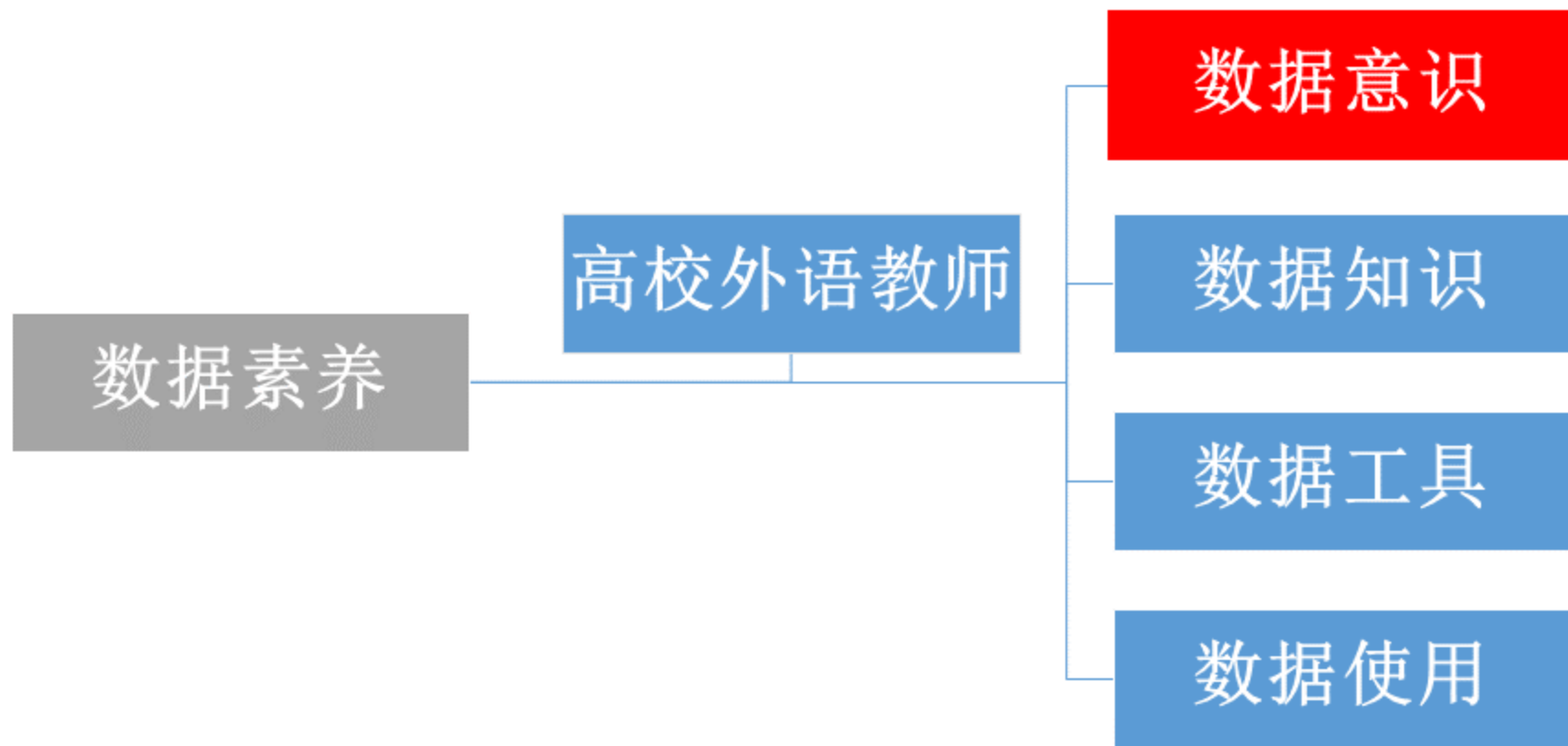




(二)

数据素养与专业发展

案例分析





选手使用的例子都很有意思!

这些例子的使用究竟如何呢?

例据: 类型、功能、质量 .....

## Assessed Levels of Second Language Speaking Proficiency: How Distinct?

<sup>1</sup>NORIKO IWASHITA, <sup>2</sup>ANNIE BROWN, <sup>3</sup>TIM McNAMARA and  
<sup>3</sup>SALLY O'HAGAN

<sup>1</sup>The University of Queensland, <sup>2</sup>Ministry of Higher Education and Scientific Research, Abu Dhabi, UAE, <sup>3</sup>The University of Melbourne

Table 2: Summary of inter-coder reliability

Category	Coding units	N	Spearman–Brown	
<i>Linguistic Resources</i>				
<i>Grammatical accuracy</i>				
Specific types of errors	1. Correct			
	article	19	0.96	
	plural	19	0.98	
	preposition	19	0.98	
	tense	19	0.99	
	third person singular	19	0.91	
	2. Error			
	article	19	0.93	
	plural	19	0.86	
	preposition	19	0.88	
	tense	19	0.71	
	third person singular	19	0.78	
	Global accuracy	Error free T-unit	19	0.98
		Error T-unit	19	0.99
<i>Grammatical complexity</i>				
T-unit complexity ratio	T-unit	20	0.91	
	Clause	20	0.94	
Dependent clause ratio	Dependent clause	20	0.99	
Verb phrase ratio	Verb phrase	20	0.98	
MLU <sup>a</sup>	Morphemes	20	1.00	
<i>Fluency</i>				
Filled pauses (um and ers)	Filled pauses	20	1.00	
Repair	Repair	20	0.98	
MLR <sup>b</sup>	Syllables	20	1.00	
	Utterances		0.98	

<sup>a</sup>MLU = mean length of utterance.

<sup>b</sup>MLR = mean length of runs.

- + 个人情况或经历
- + 个人行为与性格特征
- + 日常生活
- + 人际关系
- + 家庭、朋友与周围的人
- + 学校生活
- 语言学习
  - + 语言学习
  - + 文学阅读与艺术赏析
- + 历史与人文
- + 世界与地理
- + 兴趣与爱好
- + 文娱与体育
- + 文学阅读与艺术赏析
- + 旅游与交通
- + 假日休闲
- + 自然与环境
- + 哲人哲理
- + 社会

例据1:

I'm very glad to see that Mr. Lin Wusun on this video, because he is my idol translator. 【原文】

例据2:

volunteers, as the master of global language, English, we can play the vital part of helping people communicating with each other. For example, we can use English to know the needs of foreign athletes, and then help them. For example, we can take the foreigners all around Beijing to see the beautiful scenery and all the cultural relics of China. 【原文】

例据3:

First, we need to increase the quality of liberal arts education. In liberal arts education, the combination of knowledge is the foundation of good translation. 【原文】

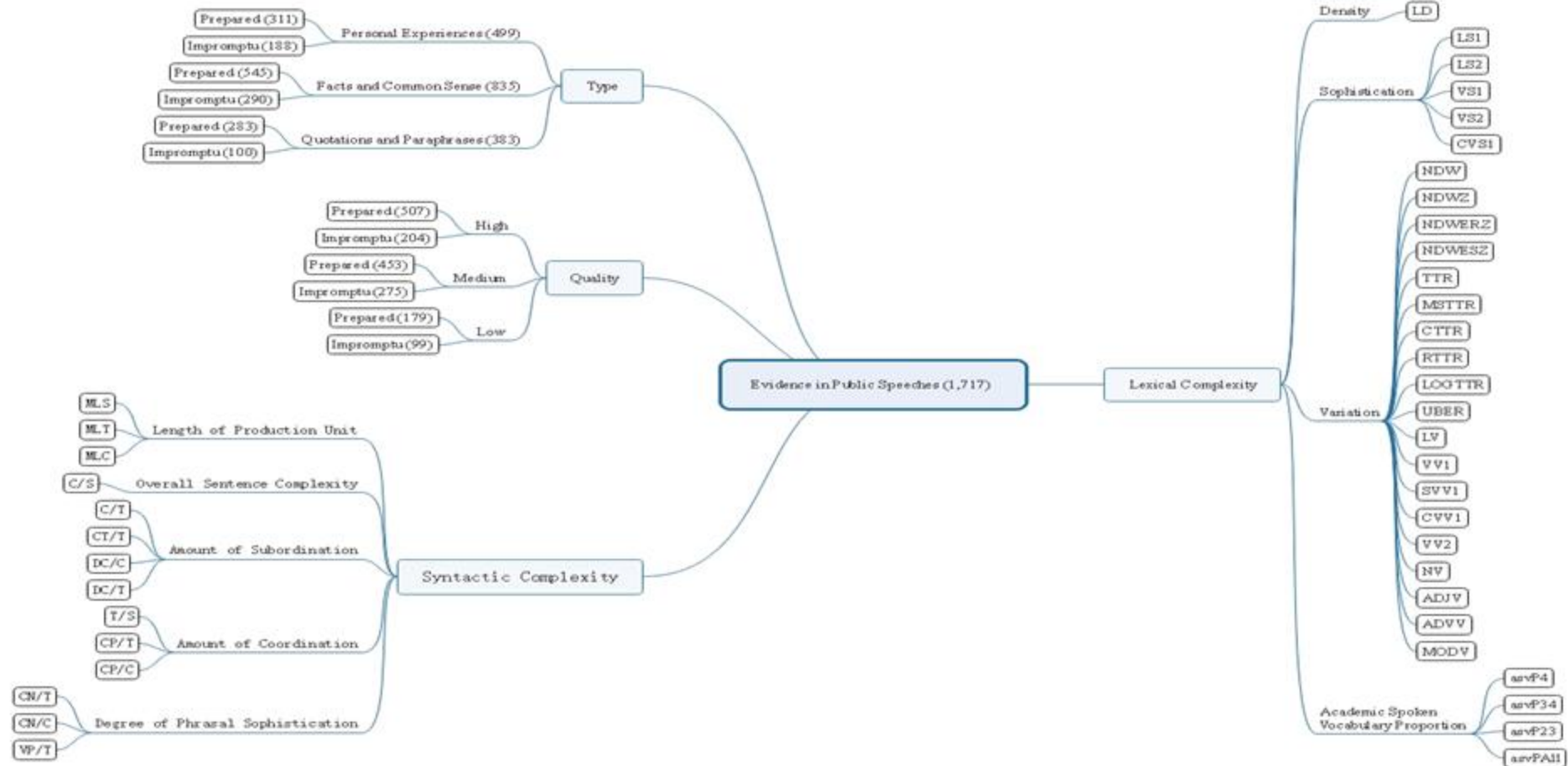
例据4:

We too often hear people complain about things. I heard people complain about their English. They said, "I cannot be a successful English speaker, because I have never been to the U.S. or the U.K." To that, we can all offer the reply. Most of the contestants sitting here have not been to the English-speaking countries, yet they can still be excellent English speakers. 【原文】

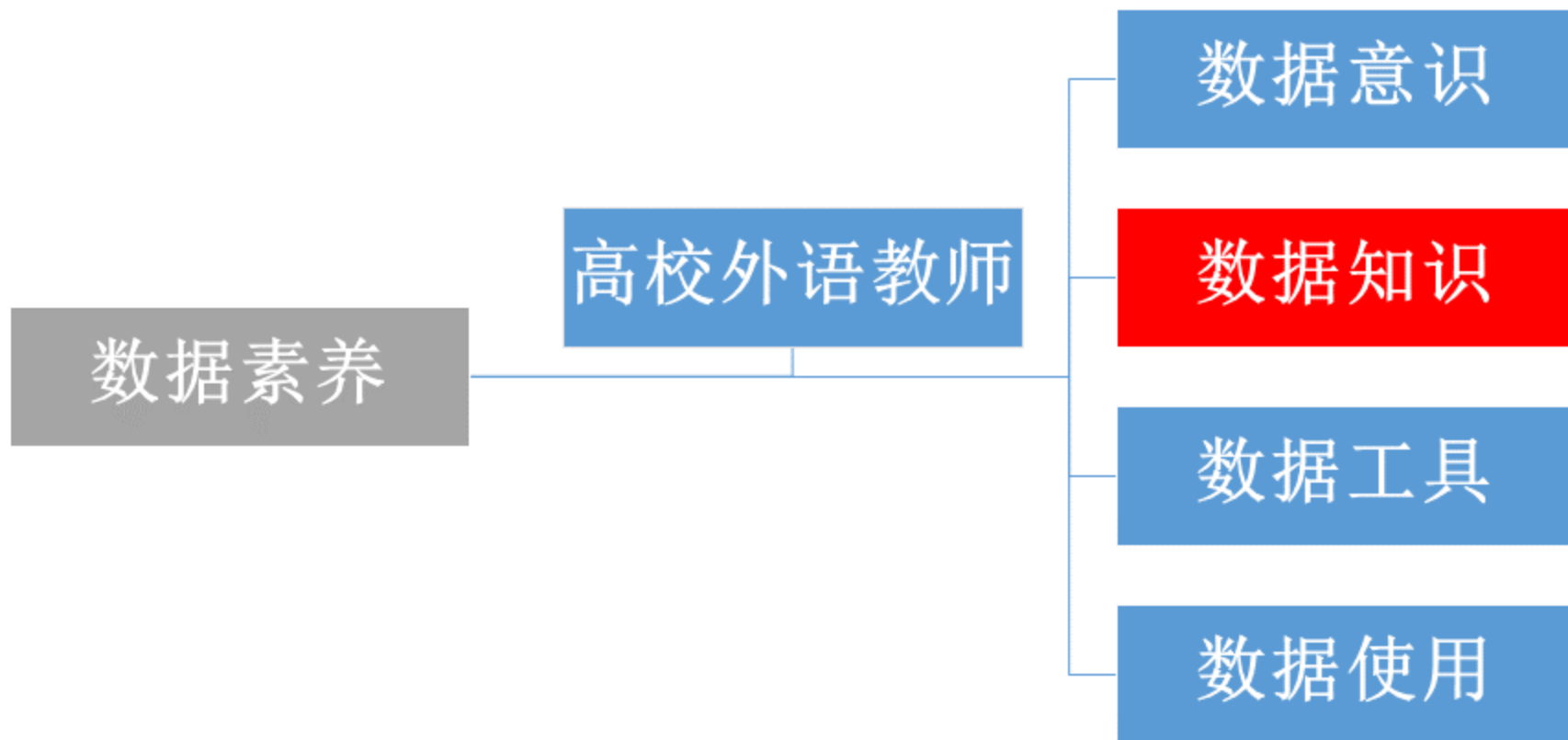
例据5:

I would like to start my speech with a question. There are thousands of different languages, spoken by people with totally different histories and backgrounds. How is it that they could be translated and share so many words that express the same meanings? 【原文】





**The Corpus of Evidence in Public Speeches**





全国统计教材编审委员会“十二五”规划教材

# 统计学： 从数据到结论

第四版

吴喜之 编著

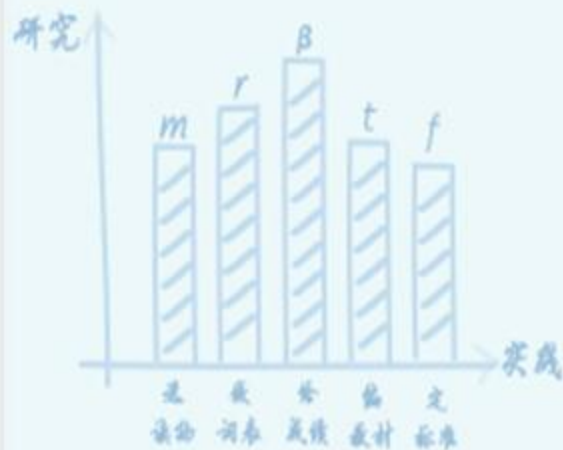
中国统计出版社  
China Statistics Press

第一章 一些基本概念 .....	1
1.1 统计是什么? .....	1
1.2 现实中的随机性和规律性,概率和机会 .....	3
1.3 变量和数据 .....	3
1.4 变量之间的关系 .....	4
1.4.1 定量变量间的关系 .....	5
1.4.2 定性变量间的关系 .....	7
1.4.3 定性和定量变量间的混和关系 .....	8
1.5 统计、计算机与统计软件 .....	8
1.6 小结 .....	11
1.7 习题 .....	11



语言、数据与研究 1

# 阿檀小倪讲量化





# 课程目录

unipus

01

研究：用数据回答问题

02

描述：大致情况如何？

03

相关：它俩有何关系？

04

回归：它们有何关系？

05

对比：它俩有何差异？

06

比较：它们有何差异？

描述：选读物 TQ

相关：做词表 AL

回归：给成绩 LT

对比：编教材 MLJ

比较：定标准 ER



描述：选读物 TQ



## Is Text Written for Children Useful for L2 Extensive Reading?

STUART WEBB AND JOHN MACALISTER

Victoria University of Wellington  
Wellington, New Zealand

The researchers completed a corpus-driven analysis of 688 texts written for children, language learners, and older readers to determine the vocabulary size necessary for comprehension and the potential to incidentally learn vocabulary through reading each text type. The comparison between texts written for different audiences may indicate their relative value for use in extensive reading programs. The results indicate that a vocabulary size of 10,000 words plus knowledge of the proper nouns and marginal words was required to know 98% of the words in both text written for children and text written for older readers. In contrast, a vocabulary size of 3,000 word families plus knowledge of the proper nouns and marginal words was necessary to know 98% of the words in text written for language learners. Repetition of words in Nation's (2006) 3rd to 14th 1,000-word lists was higher in the text written for language learners, followed by children's literature and then text written for adults. The findings indicate that the lexical load of text written for children is similar to that of text written for older readers, and that neither of these text types is as well suited as graded readers for second language extensive reading.

doi: 10.1002/tesq.70

相关：做词表 AL

Applied Linguistics 2015: 36/1: 1–22 © Oxford University Press 2013  
doi:10.1093/applin/amt018 Advance Access published on 26 August 2013

## Is There a Core General Vocabulary? Introducing the *New General Service List*

<sup>1</sup>\*VACLAV BREZINA and <sup>2</sup>DANA GABLASOVA

<sup>1</sup>Department of Linguistics and English Language, County South, Lancaster University,  
Lancaster LA1 4YL, UK

<sup>2</sup>Lancaster University, UK

\*E-mail: v.brezina@lancaster.ac.uk; d.gablasova@lancaster.ac.uk

The current study presents a *New General Service List (new-GSL)*, which is a result of robust comparison of four language corpora (*LOB*, *BNC*, *BE06*, and *EnTenTen12*) of the total size of over 12 billion running words. The four corpora were selected to represent a variety of corpus sizes and approaches to representativeness and sampling. In particular, the study investigates the lexical overlap among the corpora in the top 3,000 words based on the *average reduced frequency (ARF)*, which is a measure that takes into consideration both frequency and dispersion of lexical items. The results show that there exists a stable vocabulary core of 2,122 items (70.7%) among the four corpora. Moreover, these vocabulary items occur with comparable ranks in the individual wordlists. In producing the *new-GSL*, the core vocabulary items were combined with new items frequently occurring in the corpora representing current language use (*BE06* and *EnTenTen12*). The final product of the study, the *new-GSL*, consists of 2,494 lemmas and covers between 80.1 and 81.7 per cent of the text in the source corpora.

回归：给成绩 LT

Article

LANGUAGE  
TESTING

## Distinguishing features in scoring L2 Chinese speaking performance: How do they work?

Tan Jin and Barley Mak

The Chinese University of Hong Kong, China

Language Testing  
36(1) 23–47  
© The Author(s) 2012  
Reprints and permissions:  
sagepub.com/journalsPermissions.nav  
DOI: 10.1177/0265552212440637  
ljt.sagepub.com  
SAGE

### Abstract

For Chinese as a second language (L2 Chinese), there has been little research into distinguishing features (Fulcher, 1996; Iwashita et al., 2008) used in scoring L2 Chinese speaking performance. The study reported here investigates the relationship between the distinguishing features of L2 Chinese spoken performances and the scores awarded by raters to the performances using holistic scoring. Seven distinguishing features – representing four major categories of Pronunciation, Fluency, Vocabulary and Grammar in the L2 Chinese speaking construct – were employed. An L2 Chinese speaking test was developed to assess the overall communicative ability in L2 spoken Chinese within an academic context. Speech samples of 66 candidates on the speaking test (i.e. 9 minutes' speech length for each candidate) were analyzed in terms of the seven distinguishing features, with correlations and standard multiple regression being employed. Results showed that, first, each of the seven distinguishing features was significantly correlated to the scores, producing large or medium effect sizes; second, 79% and 77% of the variance in the scores could be explained by the distinguishing features (incorporating word tokens and word types respectively) in two regression analyses. The current study has established a link between distinguishing features and scores, contributing empirical evidence of candidate performance to the validation of assessing speaking proficiency in the L2 Chinese context.

### Keywords

Correlations, distinguishing features, scores, speaking performance, standard multiple regression

对比：编教材 MLJ

比较：定标准 ER



# The Modern Language Journal

---

## A Linguistic Analysis of Simplified and Authentic Texts

SCOTT A. CROSSLEY<sup>1</sup>, MAX M. LOUWERSE<sup>2</sup>, PHILIP M. McCARTHY<sup>3</sup> and DANIELLE S. McNAMARA<sup>4</sup>

Version of Record online: 16 FEB 2007  
DOI: 10.1111/j.1540-4781.2007.00507.x

Issue



The Modern Language Journal  
Volume 91, Issue 1, pages 15–30, Spring 2007

## Educational Researcher



3.827 Impact Factor [more »](#)

Home | Browse | Submit Paper | About | Subscribe

### Challenging the Research Base of the Common Core State Standards

*A Historical Reanalysis of Text Complexity*  
David A. Gamson, Xiaofei Lu, Sarah Anne Eckert

First Published October 1, 2013 | Research Article | [Check for updates](#)

[Download PDF](#) [Article information](#)  

#### Abstract

The widely adopted Common Core State Standards (CCSS) call for raising the level of text complexity in textbooks and reading materials used by students across all grade levels in the United States; the authors of the English Language Arts component of the CCSS build their case for higher complexity in part upon a research base they say shows a steady decline in the difficulty of student reading textbooks over the past half century. In this interdisciplinary study, we offer our own independent analysis of third- and sixth-grade reading textbooks used throughout the past century. Our data set consists of books from 117 textbook series issued by 30 publishers between 1905 and 2004, resulting in a linguistic corpus of roughly 10 million words. Contrary to previous reports, we find that text complexity has either risen or stabilized over the past half century; these findings have significant implications for the justification of the CCSS as well as for our understanding of a “decline” within American schooling more generally.

#### Keywords

curriculum, educational reform, history, reading, textbooks

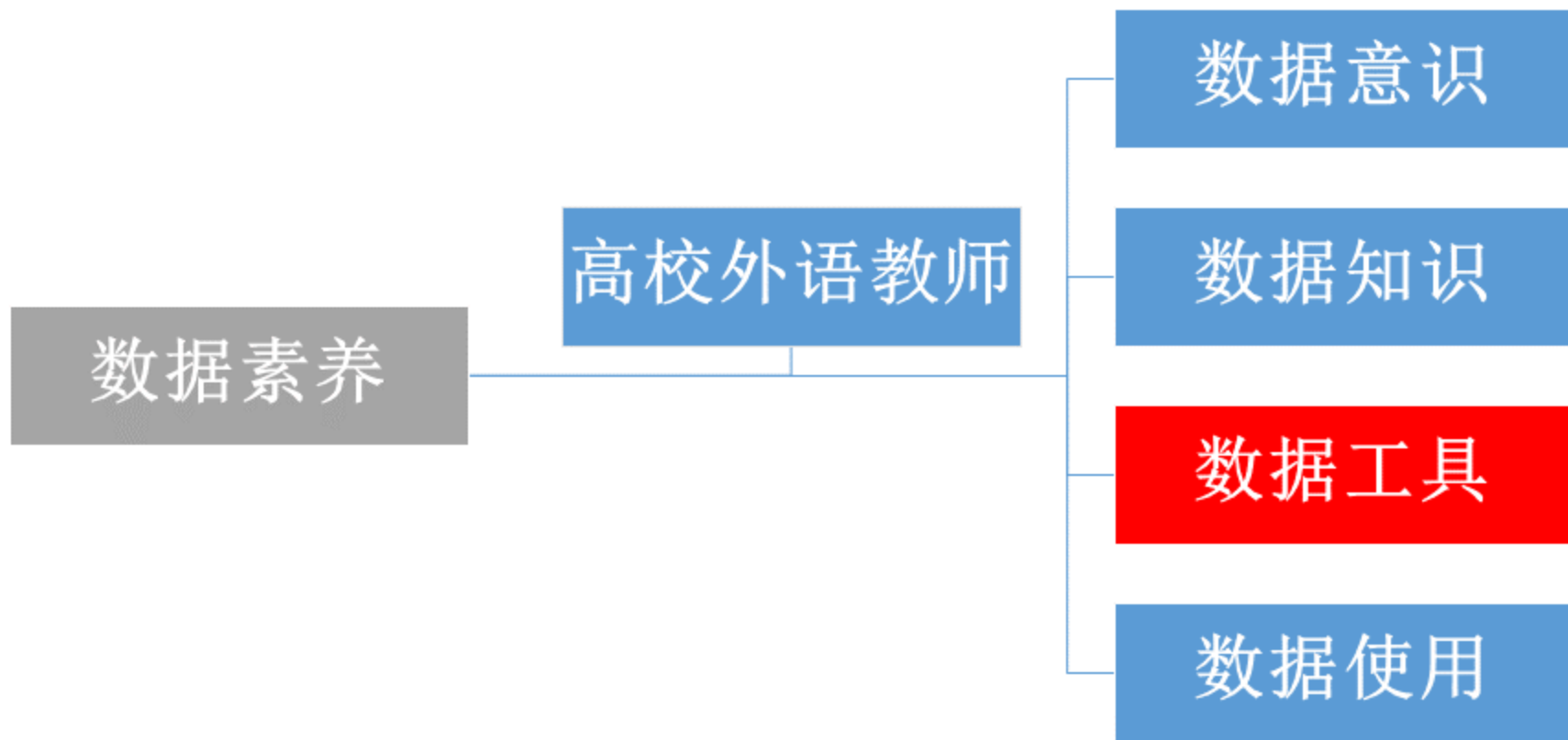
Vol 42, Issue 7, 2013

[Table of Contents](#)

- Full text +
- Supplemental material
- Figures & Tables
- Article Metrics
- Related Articles

[Cite](#) [Permissions](#) [Share](#)

[Download PDF](#)



## Contents

Language Testing Volume 29 Number 1 January 2012

### Messick Lecture, 2010

Validating score interpretations and uses

*Michael Kane*

### Responses to Kane

Validity argument for language assessment: The framework is simple...

*Carol A. Chapelle*

Grounding the argument-based framework for validating score interpretations and uses

*John W. Oller, Jr.*

Kane, validity and soundness

*Alan Davies*

### Articles

Confidence scoring of speaking performance: How does fuzziness become exact?

*Tan Jin, Barley Mak and Pei Zhou*

Note-taking quality and performance on an L2 academic listening test

*Min-Young Song*

TOEFL iBT speaking test scores as indicators of oral communicative language proficiency

*Brent Bridgeman, Donald Powers, Elizabeth Stone and Pamela Mollan*

Re-fitting for a different purpose: A case study of item writer practices in adapting source texts for a test of academic reading

*Anthony Green and Roger Hawkey*

Factor structure of the revised TOEIC® test: A multiple-sample analysis

*Yo In'nami and Rie Koizumi*

### Book reviews

Cognitive Bases of Second Language Fluency

*Norman Segalowitz, reviewed by Slobodanka Dimova*

Linguistic Relativity in SLA: Thinking for Speaking

*Z.H. Han and T. Cadierno (Eds), reviewed by Alan Davies*



## Confidence scoring of speaking performance: How does fuzziness become exact?

Tan Jin, Barley Mak Pei Zhou

First Published June 15, 2011; pp. 43–65

[Abstract](#)

[> Preview](#)



## Re-fitting for a different purpose: A case study of item writer practices in adapting source texts for a test of academic reading

Anthony Green, Roger Hawkey

First Published October 24, 2011; pp. 109–129

[Abstract](#)

[> Preview](#)



**Table 4.** Text length before and after item writer adaptations and number and types of edits made

Writer	Extent	Deletion	Consolidation	Expansion	Permutation	Move	Substitution	Insertion	Total
Jane	Paragraph length	13							13
<i>Original</i>	Sentence cluster		3		1				4
1870	Sentence	5	6	2		1			14
<i>Adapted</i>	Word/phrase	8					5	1	14
937	Total	26	9	2	1	1	5	1	45
Anne	Paragraph length	8							8
<i>Original</i>	Sentence cluster	2							2
1606	Sentence	9	1	2	1			1	14
<i>Adapted</i>	Word/phrase	6					19	5	30
946	Total	25	1	2	1		19	6	54
William	Paragraph length	2							5
<i>Original</i>	Sentence cluster	16	4						17
2518	Sentence	5	7	3	1	4		3	23
<i>Adapted</i>	Word/phrase	3				2	6	4	15
909	Total	26	11	3	1	6	6	7	60
Elisabeth	Paragraph length	6							6
<i>Original</i>	Sentence cluster	3			1				4
1592	Sentence	9	3	1	3	1			17
<i>Adapted</i>	Word/phrase	11					7	3	21
664	Total	29	3	1	4	1	7	3	48





## 英语文本指难针\*

[操作指引](#)

★参考：“语言、数据与研究”系列课程 [点击访问](#)

\*引用：全檀、李百川. (2016). “英语文本指难针”. 网址: [LanguageData.net/tester](http://LanguageData.net/tester)

文章

内容

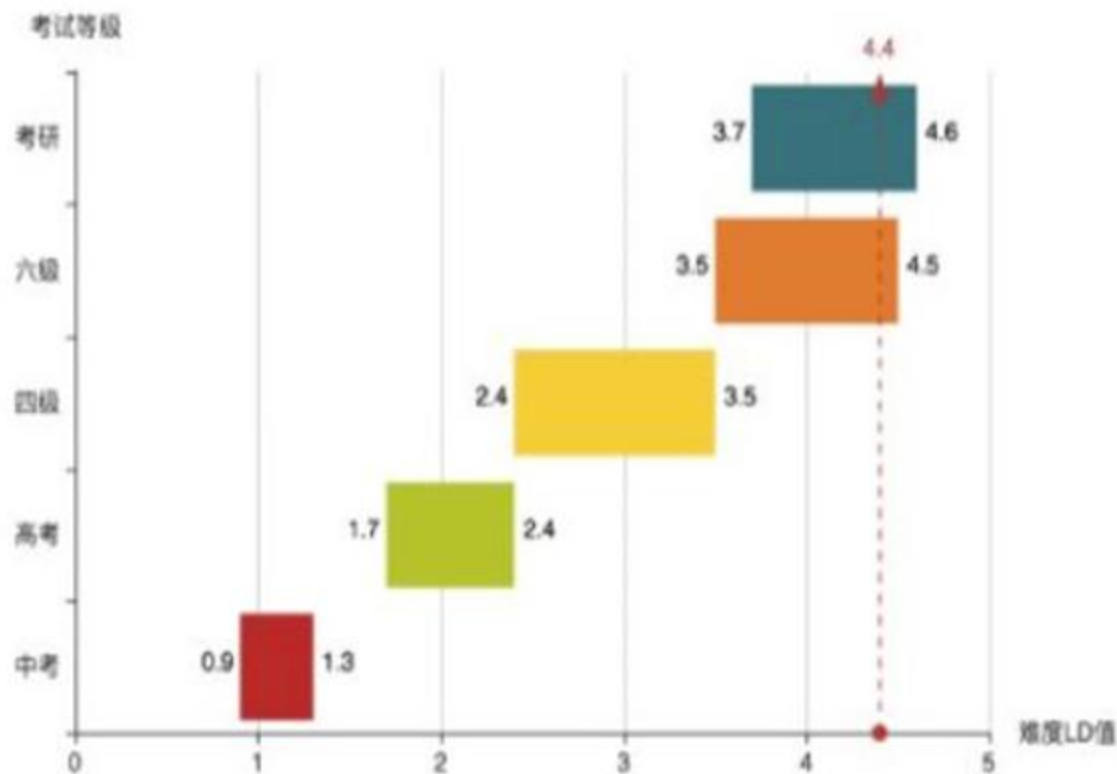
验证码



开始分析

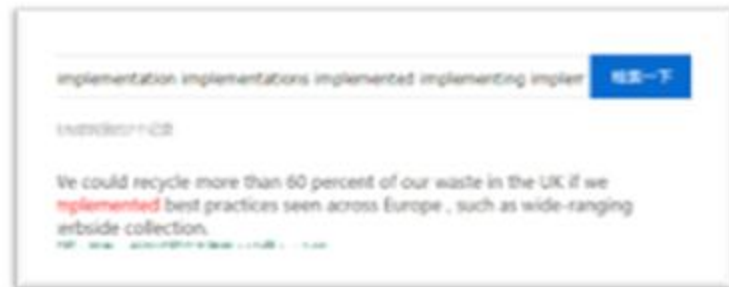
## 功能 1: 难度定级

访问网址<http://languageata.net/tester>, 进入英语文本指难针操作界面。在文本框输入需要分析的文本内容, 在验证码框中输入图示验证码, 点击“开始分析”按钮即可自动分析文本。



## 改编难度

- 中考 ● 高考 ● 四级 ● 六级 ● 考研



development (PD) course was designed and **implemented**<sup>7</sup> to a larger group of teachers. Third, in reflective practice, the PD course were analyzed to further examine the knowledge and skills that teachers can learn and benefit from data-driven text adaptation. The study also investigated the impact of the data-driven adapting actions by teachers. **Implications**<sup>12</sup> are drawn from the study to suggest human-machine **collaboration**<sup>17</sup> for producing testing materials.

oun  
释义: an article serving to equip.  
例句: the implements of religious worship.  
释义: a device used in the performance of a task.  
近义词: tool, utensil  
释义: one that serves as an instrument or tool.  
例句: the partnership agreement does not seem to be a very potent implement H. B. ...

反馈维度	原文	参考基准	改编提示
超纲词比例	15%	4%~9%	偏多

## 功能 4: 词表生成

基于标记出的超纲词, 英语文本指难针可以直接为教师生成超纲词表。仅需点击右下角“超纲词表”链接, 即可获得按英语试题语料库中出现频次(简称“基准频次”)排序的超纲词表。

排序	单词	基准频次	文中频次	覆盖率	累积覆盖率
1	archaeology	14	1	0.24	0.24
2	ecosystem	13	1	0.24	0.48
3	minimise	8	1	0.24	0.72
4	surpass	7	1	0.24	0.96
5	vocal	7	1	0.24	1.2
6	monarch	6	1	0.24	1.44
7	sovereign	6	1	0.24	1.68
8	heritage	5	1	0.24	1.92
9	reign	4	1	0.24	2.16
10	esteem	4	1	0.24	2.4
11	disrespect	4	1	0.24	2.64
12	renaissance	4	1	0.24	2.88

## 功能 7: 结果导出

在改编界面中, 教师可以在页面“原文改编”的文本框内直接进行修改, 该文本框保留了超纲词标记以方便查找。改编完成后, 点击“导出文本”, 可导出改编后的txt文本。

原文改编

"The ancient Hawaiians were astronomers," wrote Queen Liliuokalani, Hawaii's last reigning<sup>4</sup> monarch<sup>4</sup>, in 1897. Star watchers were among the most esteemed<sup>4</sup> members of Hawaiian society. Sadly, all is not well with astronomy in Hawaii today. Protests have erupted over construction of the Thirty Meter Telescope(TMT), a giant ob...

At issue is the TMT's planned location on Mauna Kea<sup>2</sup>, a d... Hawaiian islands to the heavens. But Mauna Kea<sup>2</sup> is also th... Mauna Kea's<sup>2</sup> peak rises above the bulk of our planet's de... unsurpassed<sup>2</sup> clarity.

Opposition to telescopes on Mauna Kea<sup>2</sup> is nothing new. A presence as disrespectful<sup>4</sup> for sacred land and a painful re... Some blame for the current controversy belongs to astronom... way of understanding the world. They did not always priorit... island's inhabitants. Hawaiian culture is not a relic<sup>3</sup> of the... Yet science has a cultural history, too, with roots going back... that first brought early Polynesians<sup>2</sup> to Hawaii's shores. t... telescopes on Mauna Kea<sup>2</sup> or to ban future development... his questions about who we are, where we come from, an... answering a primal<sup>6</sup> calling to know ourselves and our true... The astronomy community is making compromises to oban... visibility around the island and to avoid archaeological!

old ones will be removed at the end of their lifetimes and their sites returned to a ones will be removed at the end of their lifetimes and their sites returned to a on Mauna Kea<sup>2</sup> to embrace their cultural heritage<sup>5</sup> and to study the stars.

导出文本.txt

导出文本 清除

## 英语文本指难针\*

英语文本指难针是专门为iTEST命题员而设计的在线命题辅助工具,可实现计算机技术与人工命题的有机结合,从而帮助命题员确保命题质量、提高命题效率。

\* 金耀、李百川。(2016).“英语文本指难针”.广州:语言数据网(languagedata.net/tester).本工具研制受到外研社外研在线iTEST项目组资助。

### 操作指引

#### 功能 1: 文本定级

访问网址<http://languagedata.net/tester>后,即可进入指难针操作界面,在上方文本框中输入需要分析的文本内容,在下方验证码框中输入相应验证码,并点击“开始分析”按钮,工具即可开始自动分析。



在完成文本输入之后,工具即可对本文进行自动分析,完成难度定级并呈现分析结果。难度评估分析结果由“文本核心指标计算结果”表和“LD值使用指引图”两部分组成。“文本核心指标计算结果”表可为命题员提供原文的词汇覆盖、平均句长、最长句长、文章长度和难度LD值五项指标。其中,难度LD值是利用词汇、句子结构及句法等特征计算出的文章整体难度(数值范围0-5,数值越大表示难度越大)。

1/14

## 汉语文本指难针\*

汉语文本指难针以“国际汉语教材语料库”为数据基准,参考中华人民共和国教育部、国家语言文字工作委员会编制的《汉语国际教育用音节汉字词汇等级划分》,孔子学院总部、国家汉办编制的《国际汉语教材通用课程大纲》,专门为国际汉语文本改编提供水平定级、字词标注等多重功能。

\* 金耀、李百川。(2016).“汉语文本指难针”.广州:语言数据网(languagedata.net/editor).本工具研制受到中山大学青年教师培育项目(编号:16wkpy09)资助。

### 操作指引

#### 功能 1: 文本定级

访问网址<http://languagedata.net/editor>后,即可进入指难针操作界面,在上方文本框中输入需要分析的文本内容,在下方验证码框中输入相应验证码,并点击“开始分析”按钮,工具即可开始自动分析。



输入一篇文本后,得到了这些结果:文本的难度值(用“LD”表示)、文本等级、平均句长、最长句长、文本长度等。LD取值范围是1-4,“1”代表难度最低,“4”代表难度最高

5/14

## 英语分级指难针\*

英语分级指难针运用语言数据智能技术,以“英语教材语料库”为数据基准,为中小学校本/课外英语阅读资源进行难度分级与改编反馈。“英语分级指难针”具有文本定级、彩虹词汇、难词释义、例句查询和文本改编等功能。

\* 金耀、李百川、郭凯。(2017).“英语分级指难针”.广州:语言数据网(languagedata.net/reader).本工具研制受到广东省社科项目(编号:GD17WXXZ01)资助。

### 操作指引

#### 功能 1: 文本定级

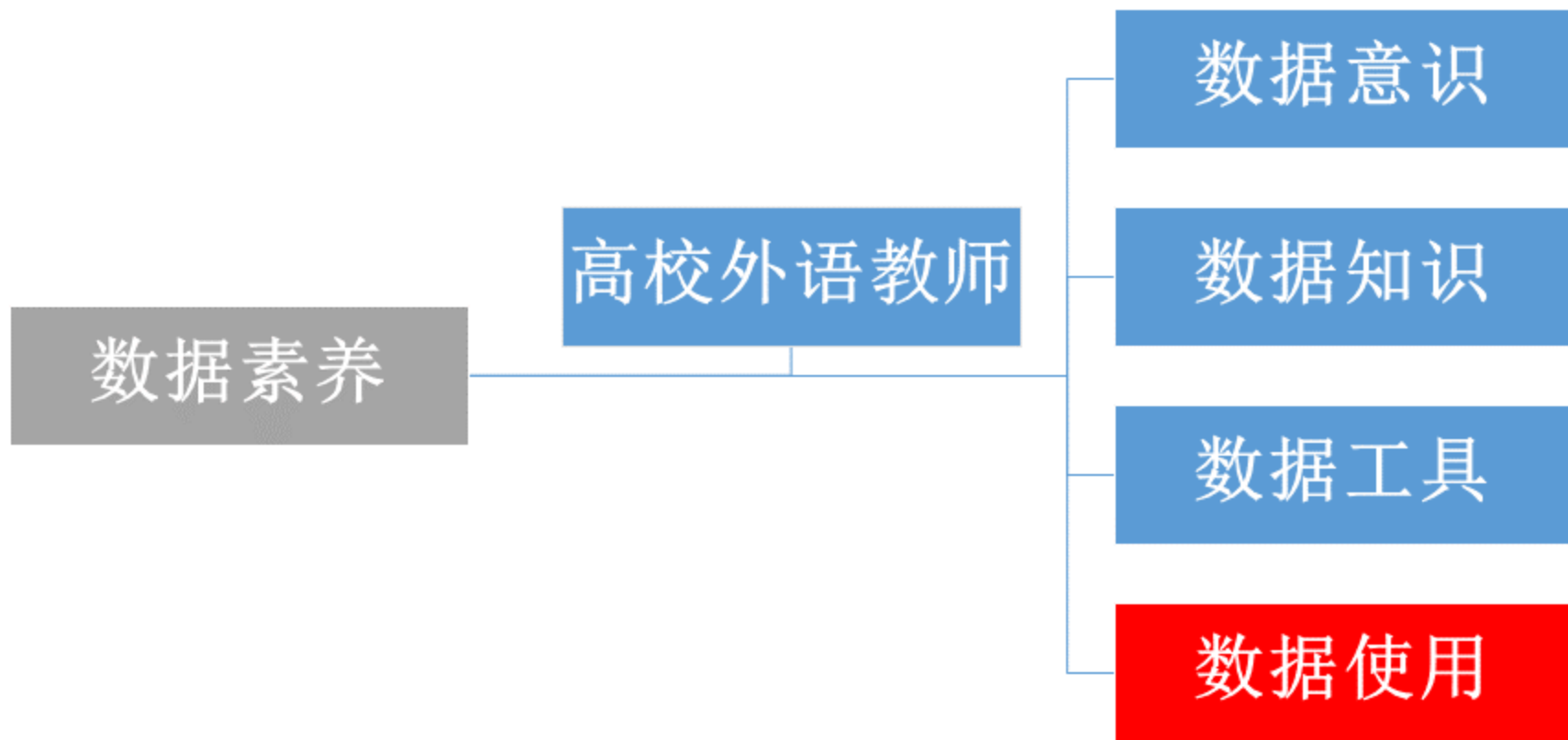
访问网址<http://languagedata.net/reader>后,即可进入指难针操作界面,在上方文本框中输入需要分析的文本内容,在下方验证码框中输入相应验证码,并点击“开始分析”按钮,工具即可开始自动分析。



这样,工具就会自动对文本进行难度分析,生成难度指数,我们可以一起看看下图中的分析结果,“核心指标计算结果表”和“LD值使用指引图”。在表中,我们可以了解文本的词汇覆盖、平均句长、最长句长、文章长度和难度LD值等指标信息;在图中,我们可以找到文本对应的难度等级。

10/14







## PreSENT

### A Basic Course for English Presentation

PreSENT是由美联英语联合语言数据网共同打造，面向我国中学生和成绩优异的小学高年级学生的英语核心演讲课程。PreSENT聚焦近10年中高考常见英语口语和写作话题，采用独家研发的“英语演讲学长说”课程模式，为学生提供“思维导图.词网句群.典型示范”三位一体的学习资源。PreSENT由美联资深外教全程讲授，重点培养公共演讲技巧，精心打造国际演讲能力。

- 思维导图——破解没有思路、缺乏条理
- 词网句群——破解词汇匮乏、句型单一
- 典型示范——破解发音不准、内容空洞

### PreSENT课堂上， 外教会教哪些公共演讲技巧？

- 如何用眼神交流
- 如何来制作PPT
- 如何用体态语言
- 如何与观众互动
- 如何用记录卡片
- 如何来回答提问

# 学术交流英语 常用论点例据说

主 编

金 檀 冯其凡 陈 静  
中山大学外国语学院



2017年12月版

## 学术交流英语：常用论点例据说

《学术交流英语：常用论点例据说》由中山大学金檀博士课程团队研制，主要面向托福雅思考生，聚焦托福雅思常考主题、采用独家研发的“常用论点例据说”的微课模式、提供“场景·论点·例据”三位一体的常用表达，全程助力学术英语听说读写。

- 1、课程主题** 基于历年托福雅思写作真题进行梳理，精选25个热门主题。
- 2、课程讲授** 基于拼图式学习特点，采用“常用论点例据说”的微课模式，由中美学长联合打造“场景·论点·例据”三大授课环节
- 3、课程素材** 基于语言数据网独家研制的托福雅思“写作真题库”，采用语言数据挖掘技术，研制托福雅思写作重点论点及例据。



# Science

## 阅读课程

« 1880s 1890s 1900s 1910s 1920s 1930s 1940s 1950s 1960s 1970s 1980s 1990s »

2010 2011 2012 2013 2014 2015 2016 2017 - -

2017

 <p>Science BIOMASS</p>	 <p>Science ANTIBIOTIC RESISTANCE</p>	 <p>Science ARCTIC CLIMATE</p>	 <p>Science TOMATO GENOME</p>	 <p>Science PRECISION MEDICINE</p>	 <p>Science EVIDENCE</p>	 <p>Science HYGIENE</p>	 <p>Science BIOMASS</p>
Jan 06 Vol 355, Iss 6320	Jan 13 Vol 355, Iss 6321	Jan 20 Vol 355, Iss 6322	Jan 27 Vol 355, Iss 6323	Feb 03 Vol 355, Iss 6324	Feb 10 Vol 355, Iss 6325	Feb 17 Vol 355, Iss 6326	Feb 24 Vol 355, Iss 6327
 <p>Science HOPE</p>	 <p>Science CANCER THERAPY</p>	 <p>Science HYBRID</p>	 <p>Science HYBRID</p>	 <p>Science WARM</p>	 <p>Science CHILD</p>	 <p>Science SPLIT</p>	 <p>Science EARTH</p>
Mar 03 Vol 355, Iss 6328	Mar 10 Vol 355, Iss 6329	Mar 17 Vol 355, Iss 6330	Mar 24 Vol 355, Iss 6331	Mar 31 Vol 355, Iss 6332	Apr 07 Vol 356, Iss 6333	Apr 14 Vol 356, Iss 6334	Apr 21 Vol 356, Iss 6335



## The social dilemma of autonomous vehicles

Jean-François Bonnefon, Azim Shariff and Iyad Rahwan

2016

Science 352 (6293), 1573-1576.  
DOI: 10.1126/science.aaf2654

## Codes of conduct in autonomous vehicles

When it becomes possible to program decision-making based on moral principles into machines, will self-interest or the public good predominate? In a series of surveys, Bonnefon *et al.* found that even though participants approve of autonomous vehicles that might sacrifice passengers to save others, respondents would prefer not to ride in such vehicles (see the Perspective by Greene). Respondents would also not approve regulations mandating self-sacrifice, and such regulations would make them less willing to buy an autonomous vehicle.  
Science, this issue p. 1573; see also p. 1514



## 文本核心指标计算结果

词汇量 (个)	40.96
平均句长	19.32
最长句长	60
文章长度	1516
单词10级	4.05
难度	25.25
可读性	91.23
六级	54.36
考研	55.20

## Altmetric



See more details

- Picked up by 184 news outlets
- Blogged by 33
- Referenced in 2 policy sources
- Tweeted by 980
- On 17 Facebook pages
- Referenced in 2 Wikipedia pages
- Mentioned in 17 Google+ posts
- Reddited by 2
- On 1 videos
- 30 readers on Mendeley
- 1 readers on CiteULike

## Readers on

30 Mendeley

## Demographic breakdown

Readers by professional status	Count	As %
Student > Bachelor	7	23%
Student > Master	7	23%
Student > Ph.D. Student	7	23%
Student > Doctoral Student	3	10%
Other	2	7%
Other	0	0%
Readers by discipline	Count	As %
Business, Management and Accounting	5	17%
Engineering	5	17%
Computer Science	4	13%
Psychology	4	13%
Unspecified	3	10%
Other	1	3%

## CASE 1

## THE SOCIAL DILEMMA OF AUTONOMOUS VEHICLE

**ethical** *adj.*

connected with beliefs and principles about what is right and wrong

**demographic** *n.*

the changing number of births, deaths, diseases, etc. in a community over a period of time; the scientific study of these changes

**trump** *v.*

to beat sth that sb. says or does by saying or doing sth even better

**disproportionate** *adj.*

too large or too small when compared with sth else

**hybrid** *n.*

something that is the product of mixing two or more different things

**subject** *v.*

to bring a country or group of people under your control, especially by using force

**median** *adj.*

being in the middle or in an intermediate position

**science-fiction** *n.*

a type of book, film/movie, etc. that is based on imagined scientific discoveries of the future, and often deals with space travel and life on other planets

**algorithm** *n.*

a set of rules that must be followed when solving a particular problem

**hypothetical** *adj.*

based on situations or ideas which are possible and imagined rather than real and true

**utilitarian** *adj.*

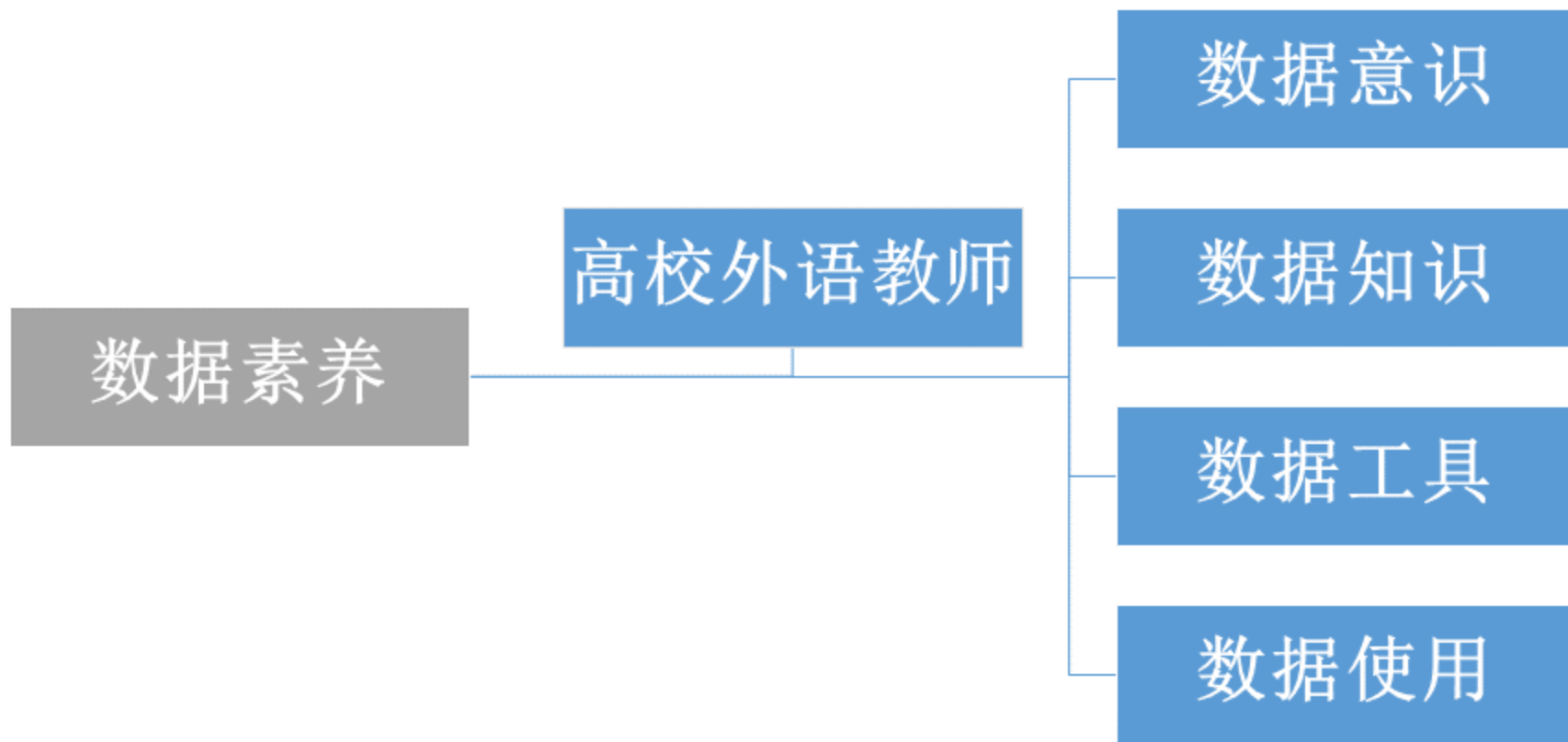
designed to be useful and practical rather than attractive

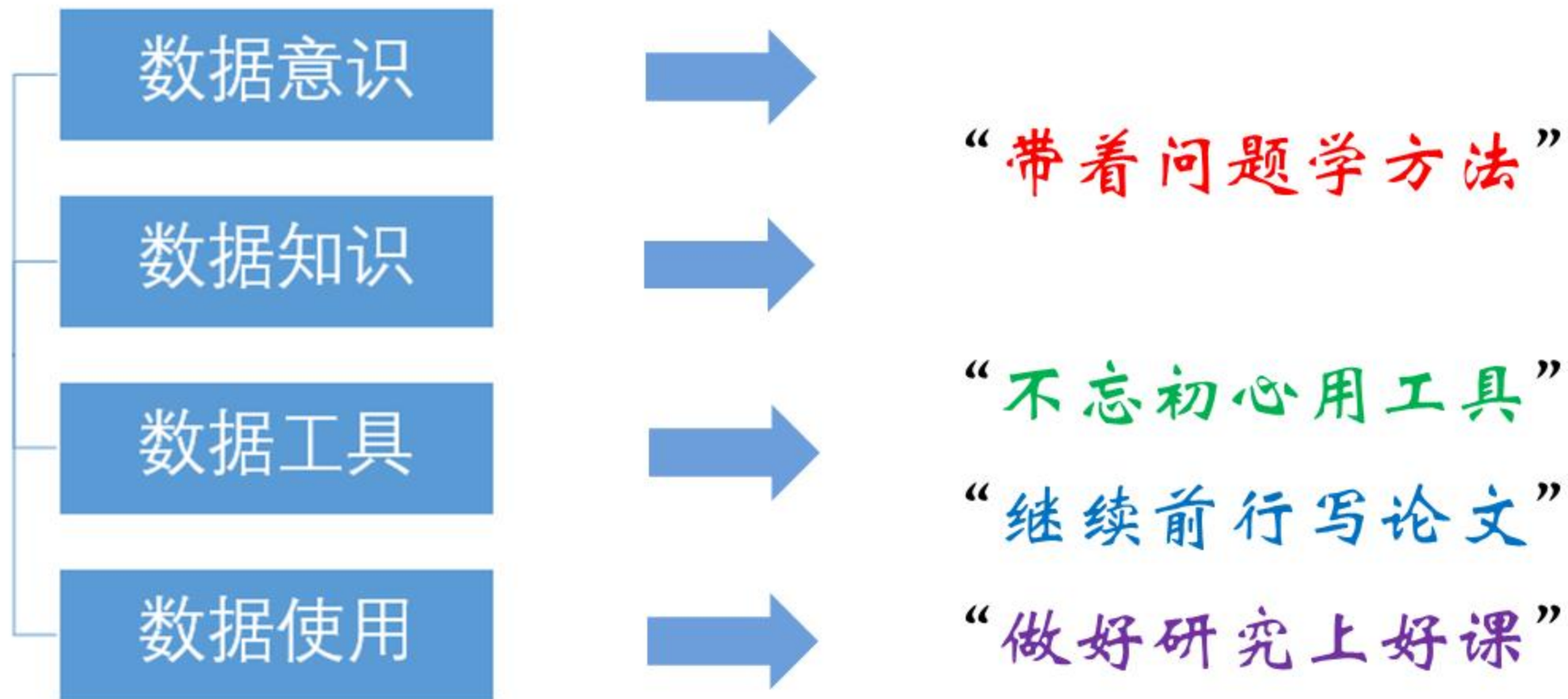
**regression** *n.*

the process of going back to an earlier or less advanced form or state



(三)  
外语教师专业发展  
实施建议





《阿檀小倪讲量化》  
“带着问题学方法”

《量化工具百宝箱》  
“不忘初心用工具”

《论文发表学者谈》  
“继续前行写论文”

《人工智能案例说》  
“做好研究上好课”

“语言、数据与研究”系列课程

主 编

中山大学 金 檀 美国宾夕法尼亚州立大学 陆小飞



语言、数据与研究1：阿檀小倪讲量化



语言、数据与研究2：量化工具百宝箱



语言、数据与研究3：论文发表学者谈



量化研究论文写作：题目亮起来，结果量出来



如何迅速找准量化研究问题



一份演讲词表为何做了十年？



如何分析作文的句法复杂度？




学术动态：如何多角度跟踪前沿？




交叉学科：如何跨学科开展合作？

和语言相伴，与数据同行，  
让我们共同成长



LANGUAGE DATA

Adaptation



语言数据素养

文本难度调控

演讲例据使用

© 2014-2018 LanguageData All Rights Reserved. 粤ICP备15030803号