

The Word List Myth

词汇表的迷思

What IELTS Reading and Listening actually test — and why I will not sell you a word list.

雅思阅读与听力真正考什么——以及我为什么不会卖给你一份词汇表

I want to start by talking myself out of a sale.

我想先把一笔生意从自己手里推出去。

Every week a student asks me the same question. *Deen, which words come up in the test? Just give me the list.* And every week I have to tell them something they do not want to hear. There is no list. Not an official one. Not one that anybody can honestly sell you.

每周都有学生问我同一个问题：“Deen 老师，考试里会出现哪些单词？把词汇表给我就行。”而我每周都必须告诉他们一个他们不想听的答案：没有这样的词汇表。没有官方版本，也没有任何人能诚实地卖给你的版本。

But I also have to tell them the other half, because the other half is true. There is a measurable core vocabulary in IELTS, and researchers have counted it carefully. So let me give you both halves properly, with every source numbered so you can check anything I say.

但我同样必须告诉他们另一半事实，因为另一半是真的。雅思确实存在一个可以量化的核心词汇，而研究者已经把它仔细统计出来了。下面我会把两半都讲清楚，并给每一个出处编号，让你可以核对我说的每一句话。

The short version

简短版本

A core vocabulary exists and has been measured. But no word list has ever been published by the organisations that own and write the test, and the paper is built so that memorised words do not turn up where you expect them. Reading and Listening are comprehension papers that use vocabulary as raw material. They are not vocabulary papers.

核心词汇确实存在，也已被量化。但拥有并编写这场考试的机构从未发布过任何词汇表；而且试卷本身的设计，就是让你背过的词不会出现在你以为它会出现的位置。阅读和听力是以词汇为原材料的理解能力考试，而不是词汇考试。

PART ONE • 第一部分

What is true: the core exists, and it has been counted

真实的一半：核心词汇存在，而且已被统计

I have been teaching English since 2005, and for most of those years this was a matter of opinion. It is not any more. Corpus linguists have taken large collections of published Cambridge IELTS practice material and counted every word in them. The findings are public and peer-reviewed, and I have listed all of them at the back so you can go and read them yourself.

我从 2005 年开始教英语，在其中大部分年份里，这个问题只能靠经验判断。现在不同了。语料库语言学家把大量公开出版的剑桥雅思练习材料集中起来，逐词统计。研究结果已公开发表并经同行评审，我把它全部列在文末，你可以自己去读。

Listening

听力

Phung and Ha analysed 239 listening transcripts drawn from 60 official practice tests — 168,370 running words in total. Coverage is measured against the BNC/COCA frequency lists. ^[1]

Phung 与 Ha 分析了取自 60 套官方练习卷的 239 份听力文本，共计 168,370 个形式词。覆盖率以 BNC/COCA 词频表为基准计算。 ^[1]

3,000 word families 3,000 个词族	95% coverage of all four sections 覆盖全部四个部分的 95%
5,000 word families 5,000 个词族	98% coverage 覆盖率达 98%
2,000 families plus the AWL 2,000 个词族 + 学术词表	95% of Sections 1-3 — which is 75% of the paper 覆盖第 1-3 部分的 95%，即整场听力的 75%
Section 4 第 4 部分	The outlier — consistently denser than 1-3 例外：词汇密度始终高于前三部分

Reading

阅读

Two teams have done this for Reading. Ha and colleagues profiled 186 passages from 42 Academic and 12 General Training tests. ^[2] Le and colleagues went back to the question with 345 passages from 115 tests. ^[3]

阅读方面有两个研究团队做过同样的工作。Ha 等人分析了来自 42 套学术类和 12 套培训类试卷的 186 篇文章。^[2] Le 等人则用 115 套试卷、345 篇文章重新审视了这个问题。^[3]

Academic passages 学术类文章	4,000–5,000 word families for 95%; 9,000 for 98% 4,000–5,000 个词族达 95%；9,000 个达 98%
General Training, Sections 1–2 培训类第 1–2 部分	3,000 word families for 95%; 6,000 for 98% 3,000 个词族达 95%；6,000 个达 98%
General Training, final passage 培训类最后一篇	Sits at Academic-level lexical demand 词汇难度与学术类相当

A word about that range. You will notice the two Reading studies do not agree. The 2022 study says 5,000 word families for 95 per cent coverage. The 2025 study says 4,000. ^[2] ^[3] Different corpora, different passage counts, different years of material.

关于这个区间。你会注意到，两项阅读研究的结论并不一致。2022 年的研究认为达到 95% 覆盖率需要 5,000 个词族，2025 年的研究认为是 4,000 个。^[2] ^[3] 语料库不同、篇目数量不同、取材年份也不同。

I could have picked whichever number sounded better and printed it as fact. I am not going to do that. Treat it as a range. And when a resource quotes you one exact number with total confidence, ask yourself where they got it.

我大可以挑一个听起来更漂亮的数字，当作定论印出来。我不会这么做。请把它当作一个区间。当某份资料信誓旦旦地给你一个精确数字时，请先问一句：这个数字从哪来的？

The Academic Word List

学术词表 (AWL)

The single most useful list in IELTS preparation was never built for IELTS. Coxhead's Academic Word List identifies 570 word families from a 3.5-million-word corpus of academic writing, and they account for roughly one word in every ten of academic text. ^[4]

雅思备考中最有用的那份词表，当初根本不是为雅思编的。Coxhead 的《学术词表》从 350 万词的学术语料中提取出 570 个词族，这些词大约占学术文本中每十个词里的一个。^[4]

It is free. It is finite. It is independently published and you can check it yourself. If you take nothing else from this document, take that.

它是免费的，是有限的，是独立发表且你可以自行核查的。如果这份文件你只记住一件事，请记住这一件。

So the honest version of the true half

因此，真实的那一半，诚实的说法是

A core vocabulary exists. It is roughly 3,000 word families for Listening and 4,000–5,000 for Academic Reading, and the 570 families of the Academic Word List are its most valuable slice. ^{[1] [2] [3] [4]}

核心词汇确实存在。听力大约需要 3,000 个词族，学术类阅读大约需要 4,000–5,000 个；而学术词表的 570 个词族，是其中最有价值的部分。^{[1] [2] [3] [4]}

PART TWO • 第二部分

What is false: nobody has published a list, and nobody can

虚假的一半：没有人发布过词汇表，也没有人能发布

The three organisations that own and write IELTS are the British Council, IDP: IELTS Australia, and Cambridge University Press & Assessment. Between them they have never published a vocabulary list for Reading or Listening. Not a frequency list. Not a ranked list. Not a required set. There is no such document, so there is nothing to memorise.

拥有并编写雅思考试的三家机构是：英国文化教育协会、IDP 教育集团（澳大利亚），以及剑桥大学出版社与考评部。三家机构从未发布过任何阅读或听力词汇表。没有词频表，没有排序表，也没有所谓的必背清单。这样的文件根本不存在，因此也无从谈起。

There is a good reason for that. Word-level data from live papers is confidential test-security material. If it were published, the test would be compromised. Every study I cite in this briefing analysed *published practice books*, not live papers — and every one of them says so in its own methodology. ^{[1] [2] [3] [5]}

这是有原因的。真题层面的词汇数据属于机密的考试安全材料，一旦公开，考试本身就失效了。本文引用的每一项研究，分析的都是公开出版的练习册，而非真题——每一项研究都在方法部分明确说明了这一点。^{[1] [2] [3] [5]}

What the test producers themselves say points the other way. The British Council's own preparation guidance tells candidates to avoid memorising long lists of complicated academic words, because it is not effective preparation and you cannot know in advance which words you will need. ^[6] IDP publishes topic-based vocabulary articles — education, environment, health, business and so on. ^[7] Those are good teaching resources written for candidates. They are not drawn from test content, and IDP does not claim they are.

考试方自己的说法，恰恰指向相反的方向。英国文化教育协会的备考指导明确建议考生不要去背诵冗长的高难度学术词表，理由是这种方式并不有效，而且你无法预知自己会用到哪些词。^[6] IDP 发布了教育、环境、健康、商业等主题词汇文章。^[7] 这些是写给考生的优质教学资源，但它们并非取自考试内容，IDP 也从未这样声称过。

I want to be fair here. There is nothing wrong with a well-made vocabulary book. I have written one myself. The problem is not the teaching. The problem is the word *official* on the cover when nobody official was ever involved.

这里我想说句公道话。一本用心编写的词汇书本身没有任何问题——我自己就写过一本。问题不在教学，而在于封面上印着“官方”二字，而官方从头到尾都没有参与。

Three things the word “official” does not mean on a cover

封面上的“官方”并不意味着这三件事

- **It does not mean published by the IELTS Partners.** Check the imprint page. If the publisher is not the British Council, IDP or Cambridge University Press & Assessment, the book has no official standing, whatever the title says.

它并不意味着由雅思主办方出版。请查看版权页。如果出版方不是英国文化教育协会、IDP 或剑桥大学出版社与考评部，那么无论书名写什么，它都没有官方地位。

- **It does not mean taken from real papers.** No commercial publisher has access to live item content. What is being sold is almost always a general frequency list, a repackaged Academic Word List, or an author's editorial judgement.

它并不意味着取自真题。没有任何商业出版方能接触到真题内容。你买到的几乎总是：一份通用词频表、一份重新包装的学术词表，或者作者个人的编辑判断。

- **It does not mean verified.** A list with no stated corpus, no stated method and no stated size is an opinion presented as data.

它并不意味着经过验证。一份没有说明语料来源、没有说明研究方法、也没有说明规模的词表，只是把观点包装成了数据。

PART THREE • 第三部分

Why a memorised list lets you down on the day

为什么背下来的词汇表在考场上会让你失望

Now I want to show you the thing that actually decides your band.

接下来，我想让你看看真正决定你分数的那个因素。

IELTS Reading and Listening are built so that the wording of the question rarely matches the wording of the passage or the recording. That is not an accident, and it is not unfair. It is the entire point. The item writer's job is to check whether you can recognise a meaning when it arrives dressed differently. ^[8]

雅思阅读和听力的设计，就是要让题目的用词尽量不与原文或录音的用词重合。这不是巧合，也不是不公平，这就是考点本身。命题人的任务，是检验你能否在同一个意思换了一身衣服之后，依然认得出来。 ^[8]

If the question says... 如果题目里写着……	...the text or recording is likely to say ……那么原文或录音里很可能是
environment	surroundings, habitat, setting, conditions
benefit	advantage, gain, positive effect, upside
reduce	decrease, cut back, scale down, curb
difficult	challenging, demanding, problematic, testing
main reason	principal cause, chief factor, what really drives it

Now picture what a candidate does under time pressure when they have memorised three thousand words but have never trained this one skill. They scan for the literal word from the question. They find it — often because it has been placed there on purpose as a distractor. They answer, and they are wrong.

现在想象一下：一个背了三千个单词、却从未训练过这项能力的考生，在时间压力下会怎么做？他会在原文里搜寻题目中的那个原词。他找到了——而那个词往往就是命题人故意放在那里的干扰项。他填了答案，于是错了。

I see this every week. It is the most common single reason a student sits at Band 6.0 and cannot understand why. It is not a vocabulary problem, and more vocabulary will not fix it.

这种情况我每周都能碰到。它是学生卡在 6.0 分、又想不明白原因的最常见原因。这不是词汇量的问题，再多背词也解决不了。

The same thing governs Listening. A speaker gives a figure and then corrects it. A lecturer describes an older theory before giving the current position. The answer almost always sits *after* the turn — *however, but actually, in fact, having said that*. If you are listening for words, you miss it. If you are listening for the shape of the argument, you catch it.

听力也是同样的道理。说话人先报一个数字，随后又修正；讲座人先介绍旧理论，再给出目前的结论。答案几乎总是出现在转折之后——however、but actually、in fact、having said that。如果你听的是单词，你会错过；如果你听的是论述的走向，你就能抓住。

“The question was never how many words you know. It is how many meanings you can still recognise when they arrive dressed differently.”

“问题从来不是你认识多少个单词，而是当同一个意思换了一身衣服出现时，你还认不认得出来。”

PART FOUR • 第四部分

What the paper is actually measuring 试卷真正测量的是什么

Four things. I have listed them in the order that candidates neglect them.

四件事。我按照考生忽视它们的程度排序。

1. Paraphrase recognition. Can you match a meaning across two different wordings, at speed, under pressure? This is the central skill in both papers.

同义改写识别。在时间和压力之下，你能否把两种不同的表达对应到同一个意思？这是两份试卷的核心能力。

2. Vocabulary depth, not breadth. Do you know a word's family, its collocations, its register and its common forms — or only its dictionary meaning? A word you recognise in one form only is a word you will miss in another.

词汇的深度，而非广度。你知道一个词的词族、搭配、语体和常见变形吗，还是只知道它的词典释义？只认得一种形式的词，换一种形式你就会错过。

3. Discourse and signposting awareness. Can you follow how a talk or an argument is organised, and hear the moment a claim is corrected, qualified or overturned?

语篇与提示词意识。你能否跟上一段讲话或一个论证的组织方式，并听出某个论断被修正、限定或推翻的那一刻？

4. Topic familiarity. The passages and recordings recycle a stable set of domains. Familiarity does not hand you the answer. It lowers what the text costs you to process, and that buys you time.

题材熟悉度。文章和录音反复使用一组相对固定的领域。熟悉并不会直接告诉你答案，但它降低了你处理文本的认知成本，而这会为你争取到时间。

Notice what is not on that list: the ability to recall words from a page.

请注意这份清单上没有的一项：从纸面上回忆单词的能力。

PART FIVE • 第五部分

What I tell my own students to do instead 我给自己学生的建议

- **Build from the Academic Word List.** 570 families is real, finite and defensible. Work through it in context, not on flashcards alone. ^[4]

从学术词表入手。570 个词族是真实的、有限的、经得起推敲的。请在语境中掌握它们，而不是只靠单词卡。 ^[4]

- **Learn families and collocations, not single words.** *Analyse, analysis, analytical, analytically, analyst.* Learn the shapes a word takes and the words it travels with.

学词族和搭配，不要只学单词。Analyse、analysis、analytical、analytically、analyst。记住一个词会变成什么形状，以及它通常和哪些词一起出现。

- **Train paraphrase directly.** Take any practice question, find the answer, then write out exactly how the question wording differed from the text wording. Do this for every question you got wrong. It is the highest-value exercise in Reading preparation and almost nobody does it.

直接训练同义改写。随便拿一道练习题，找到答案，然后把题目的用词与原文的用词之间的差异逐字写下来。每一道做错的题都这样做一遍。这是阅读备考中性价比最高的练习，而几乎没有人在做。

- **Build a signposting inventory.** Contrast, addition, example, sequence, correction, conclusion. Thirty phrases learnt properly will do more for Listening Section 4 than a thousand nouns.

建立一份提示词清单。转折、递进、举例、顺序、修正、总结。认真掌握三十个短语，对听力第 4 部分的帮助超过一千个名词。

- **Read and listen inside the recurring domains.** Environment, education, technology, health, history and archaeology, psychology, business. Not to predict the topic — to reduce what it costs you when it appears.

在反复出现的领域内阅读和听。环境、教育、科技、健康、历史与考古、心理学、商业。不是为了预测题目，而是为了降低它出现时你要付出的代价。

- **Measure coverage, not count.** The useful question is not “how many words do I know” but “how much of an unseen academic passage can I follow without stopping?” At roughly 95 per cent coverage you can read independently. Below that, you are rebuilding meaning from fragments — which is a polite way of describing guesswork. ^{[1] [2] [3]}

衡量覆盖率，而不是词汇量。真正有用的问题不是“我认识多少个词”，而是“一篇没见过的学术文章，我能不停顿地读懂多少？”覆盖率大约达到 95% 时，你才能独立阅读。低于这个水平，你就是在用碎片重建意思——那只是“猜”的一种委婉说法。 ^{[1] [2] [3]}

PART SIX • 第六部分

Before you buy anything 在你付钱之前

Five questions. Ask them of my books too.

五个问题。面对我自己的书，也请同样这么问。

- **Who published it?** If the answer is not the British Council, IDP or Cambridge University Press & Assessment, the word *official* on the cover carries no weight.

谁出版的？如果答案不是英国文化教育协会、IDP 或剑桥大学出版社与考评部，那封面上的“官方”二字就没有任何分量。

- **What corpus was it built from, and how large?** A credible resource names its source material and its method. An honest one also names its limits.

它基于什么语料库，规模多大？可信的资源会说明自己的材料来源和研究方法；诚实的资源还会说明自己的局限。

- **Does it teach word families and collocations, or single-word glosses?**

它教的是词族和搭配，还是只给一个中文释义？

- **Does it train paraphrase recognition, or only recall?**

它训练的是同义改写识别，还是只是背诵？

- **Does it promise you a band score?** Nobody can honestly make that promise. Vocabulary is one input among several, and no book controls the others.

它向你承诺分数吗？没有人能诚实地做出这种承诺。词汇只是多项因素之一，而没有任何一本书能控制其他因素。

IN SUMMARY • 小结

Where I stand

我的立场

There is a core vocabulary. It has been measured. Roughly 3,000 word families for Listening, 4,000–5,000 for Academic Reading, and the 570 families of the Academic Word List are the most valuable part of it. ^{[1] [2] [3] [4]}

核心词汇确实存在。它已被量化：听力约 3,000 个词族，学术类阅读约 4,000–5,000 个；其中最有价值的部分，是学术词表的 570 个词族。 ^{[1] [2] [3] [4]}

There is no official list, and there never has been one. What is sold as one is a general frequency list, a repackaged Academic Word List, or somebody's editorial judgement wearing a confident title. ^{[6] [7]}

没有官方词汇表，从来就没有过。那些以此为名出售的东西，不过是通用词频表、重新包装的学术词表，或者披着自信标题的个人观点。 ^{[6] [7]}

And this is not a vocabulary test. It is a comprehension test that uses vocabulary as its raw material. Prepare for the comprehension, and the vocabulary will do its work.

而且，这不是一场词汇考试。它是以词汇为原材料的理解能力考试。把理解能力练好，词汇自然会发挥它的作用。

REFERENCES • 参考文献

Sources, numbered 出处 (按编号)

Every numbered marker in this document points to an entry below. Please check them.

本文中的每一个编号标记都对应下列条目。请务必自行核对。

Peer-reviewed research

同行评审研究

- [1] Phung, D. H., & Ha, H. T. (2022). Vocabulary Demands of the IELTS Listening Test: An In-Depth Analysis. *SAGE Open*, 12(1). doi:10.1177/21582440221079934
- [2] Ha, H. T., Le, H. T., Phung, D. H., & Nguyen, S. D. (2022). Is “general” easier than “academic”? A corpus-based investigation into the two modules of IELTS reading test. *SN Social Sciences*, 2, 159. doi:10.1007/s43545-022-00461-1
- [3] Le, L. T. T., Ho, N. T. P., Trang, N. H., & Ha, H. T. (2025). Revisiting the Lexical Differences Between Academic and General Training IELTS Reading Tests. *SAGE Open*, 15(3). doi:10.1177/21582440251362269
- [4] Coxhead, A. (2000). A New Academic Word List. *TESOL Quarterly*, 34(2), 213–238.
- [5] Michell, C. (2022). Using Corpus Linguistics to Better Prepare Students for the IELTS Reading Exam. In *English Language Teaching: Theory, Research and Pedagogy*.

Test-producer guidance

考试方官方指导

- [6] British Council — IELTS preparation and vocabulary guidance. takeielts.britishcouncil.org
- [7] IDP: IELTS Australia — IELTS preparation hub and vocabulary resources. ielts.idp.com
- [8] IELTS Partners — published test format specifications for Listening and Reading. ielts.org

About the author. Deen Cox is an IELTS Specialist, author and curriculum developer based in Beijing. He has taught English since 2005 and is the author of the Deen Cox IELTS Series, reviewed and endorsed by Mike Wattie, a former IELTS examiner. He is also co-author, with Wattie, of four bilingual Chinese-English IELTS textbooks. deencoxielts.com

关于作者。 Deen Cox 是驻北京的认证雅思专家、作者与课程开发者。他自 2005 年起从事英语教学，是《Deen Cox 雅思系列》的作者；该系列由前雅思考官 Mike Wattie 审校并推荐。他另与 Wattie 合著四册中英双语雅思教材。deencoxielts.com

IELTS is a registered trade mark of the British Council, IDP: IELTS Australia Pty Ltd and Cambridge University Press & Assessment. This document is independent commentary. It is not affiliated with, endorsed by, licensed by or produced in association with the IELTS Partners, and no such association is claimed or implied. Research findings are attributed to their published authors and describe analyses of commercially published practice material, not live test papers.

IELTS 为英国文化教育协会、IDP Education Australia Pty Ltd 及剑桥大学出版社与考评部的注册商标。本文为独立评述，与雅思主办方无隐属、背书、授权或合作关系，也不作此类暗示。文中研究结论均归属其发表作者，且均针对公开出版的练习材料，而非真题。

© Deen Cox. This briefing may be shared and quoted freely with attribution.

© Deen Cox. 本文可在注明出处的前提下自由分享与引用。