

國家教育研究院

National Academy for Educational Research

〈臺灣華語文能力基準—華語文能力第一套標準〉

Taiwan Benchmarks for the Chinese Language—The  
First Set of Standards for Chinese Language  
Proficiency

持續提升華語文教育的品質及效能，是教育部重要的政策，因此請國家教育研究院研發華語文能力指標及分級標準，以作為華語文教學、教材設計及測驗評量的依據。學生學習是逐漸由淺入深、加廣及加深。華語文能力指標提供學習者在各學習階段應知和應能為何，而分級標準提供學習者在各學習階段應學習的漢字、詞語及語法點。

It is Ministry of Education's major policy to continuously to improve the quality and effectiveness of Chinese language education. Therefore, the National Academy for Educational Research was assigned to research and develop Chinese language proficiency level descriptors and grading scales to serve as a basis for Chinese language teaching, teaching materials design, and testing and evaluation. Students learn from the basics and gradually broaden and deepen their knowledge. The Chinese language proficiency level descriptors provide what students should know and be able to do at each stage of learning, and the grading scales provides the Chinese characters, words and grammar that students should pick up at each stage of learning.

國家教育研究院邀請學者專家成立研究團隊，除了研究團隊人員研究外，共召開超過百場各類學者專家諮詢會議，歷經 6 年才發



展深具實證基礎的華語文能力指標及符合語言使用情境及學習者能力的漢字、詞語及語法點分級標準。

The National Academy for Educational Research invited scholars and experts to form a research team. In addition to the research team's studies, more than 100 consultation meetings were held with scholars and experts in various aspects. It took more than six years to develop the evidence-based Chinese language proficiency level descriptors and grading scales for Chinese characters, words, and grammar that meet the context of language use and students' abilities.

研究團隊分析並歸納歐洲語言共同參考架構 (CEFR)、美國外語教學協會指標 (ACTFL)、歐洲漢語能力基準 (EBCL)、漢語水平等級標準與語法大綱、新漢語水平考試 (新 HSK)、國際漢語能力標準等國際有關華語文能力指標發展之文獻，建構「華語文能力指標訂定架構」，架構包含內容標準與表現標準，內容標準陳述學習者應知和應能為何 (what students should know and be able to do)，表現標準根據內容標準界定學習者表現應該多好才達到標準 (how good is good enough)。表現標準的敘寫表現分成若干標準，描述各標準應展現的表現品質或精熟程度。

The research team analysed and summarized international literature on the development of Chinese language proficiency level descriptors, including Common European Framework of Reference for Language (CEFR), American Council on the Teaching of Foreign Language (ACTFL), European Benchmarks for the Chinese Language (EBCL), Chinese Language Level Grading Scales and Grammatical Framework, New Chinese Proficiency Test (New HSK), International Chinese Language Proficiency Level Standards, etc., and constructed the “Chinese Language Proficiency Level Descriptors Setting Framework,” which consists of content standards and performance standards. The content standards state what students should know and what they should be able to do. The performance standards define how good is good enough for students to perform according to the content standards. The narrative performance of a performance standard is divided

into criteria that describe the quality or level of sophistication of performance that each standard should exhibit.

能力指標的架構包括四向度：

The framework of proficiency level descriptors includes four dimensions:

一、「場域」，包括：「個人」、「公共」、「學校與教育」、「工作與職場」。

I. “Situation,” including: “personal,” “public,” “school and education,” “job and workplace”.

二、「溝通模式」，包括：「接受」、「產出」與「互動」。

II. “Communication mode,” including “acceptance,” “output” and “interaction”.

三、「文本形式」，包括：「字詞」、「句子」、「段落」與「篇章」。

III. “Text form,” including: “words and phrases,” “sentences,” “paragraphs” and “chapters”.

四、「表現標準」，包括：「淺層理解」與「深層理解」。

IV. “Performance standards,” including: “basic understanding” and “deep understanding”.

能力指標訂為三等七級：

Proficiency level descriptors are set at 3 Grades and 7 Levels:

「三等」分別為「基礎」、「進階」、「精熟」。

“The 3 Grades” are “Basic,” “Advanced,” and “Proficient”.

「七級」則為「1級」、「2級」、「3級」、「4級」、「5級」、「6級」、「7級」。

“The 7 Levels” are “Level 1,” “Level 2,” “Level 3,” “Level 4,” “Level 5,” “Level 6,” and “Level 7”.

1 至 3 級為基礎，4 至 5 級為進階，6 至 7 級為精熟。至於翻譯能力指標設定以華語為第二語言學習者譯為目的語之能力，考量翻譯需綜合多元能力，故將等級設定自第 3 級起至第 7 級。

Levels 1 to 3 are basic levels, Levels 4 to 5 are advanced levels, and Levels 6 to 7 are proficient levels. As for the translation proficiency level, it is set from Level 3 to Level 7 for students of Chinese as a second language to translate into the target language, considering that translation requires a combination of multiple skills.

能力指標研擬完成後，研究團隊於 106 年 3 月召開會議，徵詢華語教學系所、華語文教學中心、華語文學術社團共計 35 個單位，其後並召開多次會議，針對指標內容逐條討論修訂，並於 107 年進行華語文能力指標與 ACTFL、CEFR 語言能力指標的等級對應研究。對應過程中，與會專家提出指標內容在文字上的修訂建議，因此計畫人員於 108 年 5 月再次公開徵求能力指標的修訂意見，於 7 月再次進行專家諮詢，並完成修訂。

After completing the research and development of the proficiency level descriptors, the research team held a meeting in March 2017 to consult with Chinese language teaching departments and institutes, Chinese language teaching centers, and Chinese language academic associations, totaling 35 agencies and subsequently held several meetings to discuss and revise the descriptors one by one. In 2018, the research team conducted a study on the mapping of Chinese language proficiency level descriptors to ACTFL and CEFR language proficiency level descriptors. In the process of mapping, experts suggested textual revisions to the descriptors at the meeting, so the project staff sought comments on revisions to the proficiency level descriptors again in May 2019, conducted another expert consultation in July, and completed the revisions.

分級標準的「漢字」收錄原則為「學習者必須達到識讀程度並

掌握字的書寫」，第 1 級至第 7 級共收錄 3,100 字。

The principle for the inclusion of “Chinese characters” in the grading scales is “students must be able to read and write the characters,” and 3,100 characters are included in Levels 1 to 7.

「詞語」收錄原則為「以溝通使用為主」，第 1 級至第 7 級共收錄 14,425 詞。

The principle for the inclusion of “words” is “mainly for communication,” and 14,425 words are included in Levels 1 to 7.

「語法點」收錄原則為「基本」和「常用」的教學語法，第 1 級至第 5 級共收錄 496 個。

The principle for the inclusion of “grammar” is “basic” and “common” teaching grammar, and 496 grammar usage are included in Levels 1 to 5.

國家教育研究院並已建置漢字分級標準檢索系統、詞語分級標準檢索系統和語法點分級標準檢索系統，網址為 <https://coct.naer.edu.tw/standsys/>，以方便使用者使用查詢。

The National Academy for Educational Research has established the Chinese character grading scales retrieval system, the word grading scales retrieval system, and the grammar grading scales retrieval system (<https://coct.naer.edu.tw/standsys/>) for the convenience of users.

國家教育研究院已將上述成果結合其它研究成果建置國家教育研究院華語文語料庫與能力基準整合應用系統（網址為 <https://coct.naer.edu.tw/>），可作為華語文教師教材編撰、教學和評量、華語文學習者自主學習、華語文教材編者編撰教材及國家華語測驗推動工作委員會「華語文能力測驗」(Test of Chinese as Foreign Language) 等應用。

The National Academy for Educational Research has combined the aforementioned results with other research findings to establish the “Corpus of

Contemporary Taiwanese Mandarin (COCT) (<https://coct.naer.edu.tw/>),” which can be used for Chinese language teachers’ materials development, teaching and evaluation, Chinese language students’ self-directed learning, Chinese language textbook writers’ materials development, and the Test of Chinese as Foreign Language of the Steering Committee for the Test Of Proficiency-Huayu.

華語文標準體系各級能力內容，摘述如下：

**The content of the Chinese language standards system at each proficiency level is summarized below.**

### **第 1 級 Level 1**

能理解並運用與個人生活相關的簡單字詞。

學習者必須達到識讀程度並掌握 246 個漢字的書寫，396 個詞語的溝通使用與 15 個基本和常用語法點的學習。

Able to understand and use simple words and phrases relating to one’s daily life. Students must be able to read, write 246 Chinese characters, and use 396 words and 15 basic and common grammar to communicate.

### **第 2 級 Level 2**

能理解與個人及日常生活有關的字詞、短語及短句。能運用簡單字詞與句子與他人交流。

學習者必須達到識讀程度並掌握 504 個漢字的書寫，798 個詞語的溝通使用與 107 個基本和常用語法點的學習。

Able to understand simple words, phrases and short sentences having to do with one’s daily life. Able to converse with others using simple words, phrases and sentences.

Students must be able to read, write 504 Chinese characters, and use 798 words and 107 basic and common grammar to communicate.

### 第 3 級 Level 3

能理解公共、學校及職場等周遭場域的簡短對話、討論與指示。能以簡單完整的句子敘述經歷，與人進行對話及簡單討論。

學習者必須達到識讀程度並掌握 801 個漢字的書寫，1,254 個詞語的溝通使用與 241 個基本和常用語法點的學習。

Able to understand short conversations, discussions and instructions, in public settings, schools, and the workplace. Able to use short and complete sentence to describe experiences and engage in conversations and simple discussions with others.

Students must be able to read, write 801 Chinese characters, and use 1,254 words and 241 basic and common grammar to communicate.

### 第 4 級 Level 4

面對周遭場域，能理解較長段落、熟悉主題的對話、演說、討論及視聽媒體的內容。能以連貫句群表達情感與觀點，進行討論與演說。學習者必須達到識讀程度並掌握 1,300 個漢字的書寫，2,669 個詞語的溝通使用與 390 個基本和常用語法點的學習。

In various situations, able to understand lengthy discourses on familiar topics, whether the discourse is in the form of a conversation, speech, discussion, or audiovisual media content. Able to express one's feelings, state one's viewpoint, engage in discussions, and give speeches, using connected sentence clusters.

Students must be able to read, write 1,300 Chinese characters, and use 2,669 words and 390 basic and common grammar to communicate.

### 第 5 級 Level 5

不論場域，能理解多樣主題篇章所表達的觀點。能以流暢的段落及有組織的篇章表達觀點、摘要、評論或創作。

學習者必須達到識讀程度並掌握 1,900 個漢字的書寫，5,288 個詞語的溝通使用與 496 個基本和常用語法點的學習。

Regardless of the situation, able to understand the viewpoint expressed in chapters on a wide variety of themes. Able to present viewpoints, summaries, observations, or creative writing in fluent paragraphs and organized chapters.

Students must be able to read, write 1,900 Chinese characters, and use 5,288 words and 496 basic and common grammar to communicate.

## 第 6 級 Level 6

不論場域，能理解不同形式、多樣主題篇章的觀點及隱含的意義。能以書面語及結構完整的篇章，精確表達觀點、摘要、評論或創作。學習者必須達到識讀程度並掌握 2,500 個漢字的書寫，9,432 個詞語的溝通使用與至少 496 個基本和常用語法點的學習。

Regardless of the situation, able to understand the viewpoints and connotations of chapters that take many different forms and cover a wide variety of topics. Able to present viewpoints, summaries, observations or creative writing with precision in written language and in well-structured chapters.

Students must be able to read, write 2,500 Chinese characters, and use 9,432 words and 496 basic and common grammar to communicate.

## 第 7 級 Level 7

能理解複雜主題內容與專業領域的觀點。能精準表達觀點、評論及創作。

學習者必須達到識讀程度並掌握 3,100 個漢字的書寫，14,425 個詞語的溝通使用與至少 496 個基本和常用語法點的學習。

Able to understand the content of complex subject matter and the viewpoint of professional fields. Able to present viewpoints, summaries, observations or creative writing with precision.

Students must be able to read, write 3,100 Chinese characters, and use 14,425 words and 496 basic and common grammar to communicate.



國家教育研究院

National Academy for Educational Research

華語文語料庫與能力基準整合應用系統

COCT & TBCL

## 一、教學實務所面臨的問題 Problems in teaching practice

臺灣在華語文教學雖然頗具口碑，惟缺乏標準體系、大型語料庫及應用系統，影響教學、教材設計及測驗評量的效能，說明如下：

Although Taiwan has a good reputation in Chinese language teaching, the lack of a standards system, a large corpus and an application system, affects the effectiveness of teaching, teaching material design, and testing and evaluation, explained as below.

### (一) 華語文教學缺乏標準體系

Chinese language teaching lacks a standards system

華語文標準體系包括能力指標及漢字、詞語及語法點的分級標準。分析臺灣四部最重要的華語文教材中，共同收錄的生詞僅占 9.8%；華語地區四套最常用的教材，共同收錄的語法點也僅為 13%；多份華語文字表，共同收錄的漢字差異亦很大。顯示因為缺乏華語文標準體系，教材編撰者編寫教材時，多依主觀經驗選用漢字、詞語及語法點。作為教學及測驗評量參考使用的教材差異如此大，教學及測驗評量亦有類似情形。

The Chinese language standards system includes proficiency level descriptors and grading scales for Chinese characters, words and grammar. An analysis of the four most important Chinese language teaching materials in Taiwan shows that only 9.8% of the vocabulary words are the same; in the four most commonly used

teaching materials in the Chinese-speaking regions, only 13% of the grammar are the same; and the Chinese characters that are common in several Chinese language lists vary greatly. It shows that the lack of a Chinese language standards system has led teaching material writers to include Chinese characters, words, and grammar based on their subjective experiences. With such a wide variation in the teaching materials used as reference for teaching and evaluation, there is a similar situation for teaching and testing and evaluation.

就如中小學教學需要課程綱要，華語文教學亦需要標準體系做為課程指引，以提供學習者在各學習階段應知和應能為何，學習才能由淺入深、循序漸進有系統的進行，確保並提升學生的學習成效。

Just as primary and secondary schools need a curriculum syllabus, Chinese language teaching needs a standards system as a curriculum guideline to provide students with what they should know and be able to do at each stage of learning, so that learning can proceed in a systematic and progressive manner from basic to advanced levels, ensuring and enhancing students' learning effectiveness.

## (二) 華語文標準體系建置缺乏大型華語文語料庫及應用系統

### Chinese language teaching lacks a large corpus and an application system

華語文標準體系建置若能依據語言實際使用情境，則更能符合學習的需要。語料庫建置大量實際使用的語言資料，是最好的選擇。語料庫經由應用系統分析，統計漢字、詞語、語法點及搭配詞的真實使用語境和頻率，可以應用在建置華語文標準體系；也可以提供詞義分析及例句編寫等應用在華語文教學。然而，2013年當時國內缺乏大型華語文語料庫，也無應用系統。

The Chinese language standards system can better meet the needs of learning if it is established on the actual context of language use. A corpus with a large amount of language data actually in use is the

best choice. Through the analysis and statistic by the application system of the context and frequency of actual use of Chinese characters, words, grammar, and collocations, the corpus can be used to build a Chinese language standards system. The corpus can also be applied to Chinese language teaching by providing meaning analysis and example sentence compilation. However, in 2013, there was no large-scale Chinese language corpus and no application system in Taiwan.

然而，更重要的是，即使應用語料庫完成華語文漢字、詞語及語法點分級標準，而沒有線上查詢系統，及未與語料庫及應用系統整合，都將影響其在華語文教學、教材設計及測驗評量的有效應用。

More importantly, even if the corpus is applied to complete the grading scales of Chinese characters, words, and grammar, the lack of an online query system and the lack of integration with the corpus and application system will affect its effective application in Chinese language teaching, teaching material design, and testing and evaluation.

## 二、發展華語文能力基準與應用系統

### Development of Chinese language proficiency benchmarks and application systems

#### (一) 華語文語料庫及臺灣華語文能力基準 COCT & TBCL

為了建置華語文語料庫及研發具實證基礎的華語文能力指標和漢字、詞語及語法點分級標準，國家教育研究院（以下稱本院）邀請學者專家成立諮詢團隊並組成語料庫組、能力指標組、漢字組、詞語組及語法組等 5 個研究團隊進行研發建置；除了研究團隊人員研究外，也召開超過百場各類學者專家諮詢會議，其中包括全面徵詢華語教學系所、華語文教學中心及華語文學術社團。參與研發諮詢專家之多及廣，在華語文教學界可謂空前；歷經 6 年完成臺灣第一套華語文能力基準及建置最新且最大的華語文語料庫及多套應用系統。

In order to build a Chinese language corpus and develop evidence-based Chinese proficiency level descriptors and grading scales for Chinese characters, words, and grammar, The National Academy for Educational Research (hereinafter referred to as “the Academy”) invited scholars and experts to form a consultation team of five research groups, including the corpus group, proficiency level descriptors group, Chinese characters group, words group, and grammar group, to conduct research and development. In addition to the research team’s studies, more than 100 consultation meetings were held with a variety of scholars and experts, including comprehensive consultations with Chinese language teaching institutes, Chinese language teaching centers, and Chinese language academic associations. The number and spread of experts involved in research and development is unprecedented in the field of Chinese language education. It took six years to complete the first set of Chinese language proficiency benchmarks in Taiwan and to build the latest and largest Chinese language corpus and several application systems.

在語料庫建置，本院已建置書面語語料庫 4 億 4,041 萬字（含新聞語料共 37 億 7,178 萬字）、口語語料 5,862 萬字、中介語語料 176 萬 4 千字、華英雙語語料 1,300 萬字，以及〈通用詞頻表〉。並配合能力基準的應用，開發多套應用系統。

In terms of corpus development, the Academy has built a written language corpus of 440.41 million words (including 3,771.78 million words from news articles), a spoken language corpus of 58.62 million words, an interlanguage corpus of 1,764,000 words, a Chinese-English bilingual corpus of 13 million words, and a common words frequency list. The Academy has also developed multiple application systems in concert with the application of proficiency benchmarks.

在標準體系建置，本院已完成聽、說、讀、寫、譯三等七級能力指標，三等七級 3,100 個漢字字表及 14,425 個詞的詞語表、73 個類詞綴；並完成第 1 級至第 5 級 496 個語法點及例句編寫，第 1 至

第 3\*級 1,265 個中文基礎詞彙內容編寫，及詞語表第 1 至第 5 級共 5,322 個詞的情境分類。在與國際指標接軌，完成與美國外語教學委員會（American Council on the Teaching of Foreign Language, ACTFL）及國家華語測驗推動工作委員會之「華語文能力測驗」（Test of Chinese as Foreign Language, TOCFL）（依據歐洲共同語文參考架構指標，Common European Framework of Reference for Languages, CEFR）能力指標的等級對應。

In the development of the “The Chinese language standards system,” the Academy has completed the 3 Grades and 7 Levels of proficiency level descriptors for listening, speaking, reading, writing, and translating with the list of 3,100 Chinese characters, 14,425 words, and 73 affixes; also completed the compilation of 496 grammar and example sentences from Level 1 to Level 5, and 1,265 basic Chinese vocabularies from Level 1 to Level 3\*, as well as the context classification of 5,322 words from Level 1 to Level 5 of the word list. For the alignment with the international descriptors, the Academy has completed the mapping of proficiency level descriptors to the American Council on the Teaching of Foreign Language (ACTFL) and the Test of Chinese as Foreign Language (TOCFL) of the Steering Committee for the Test Of Proficiency-Huayu (based on the Common European Framework of Reference for Languages (CEFR)).

華語文標準體系是依據教育部華語文八年計畫所研發，因此正式名稱參考歐洲漢語能力基準（European Benchmarks for the Chinese Language, EBCL），訂為臺灣華語文能力基準，英文為 Taiwan Benchmarks for the Chinese Language（簡稱 TBCL）。從 TBCL 與 TOCFL 對應圖，發現在聽力、閱讀及寫作，TBCL 的第 1 級及第 2 級都在 TOCFL 入門級（A1）以下；在口說，TBCL 的第 1 級在 TOCFL 入門級（A1）以下；顯示 TBCL 確實達成華語文專家學者深切期待的建構具華語文特性之能力基準的目標。教育部「華語文能力測驗」於 2021 年在成績單加註 TBCL 等級。

The Chinese language standards system was developed in accordance with the Ministry of Education's eight-year plan for the Chinese language, so the official name is set as the Taiwanese Benchmarks for the Chinese Language (TBCL), with reference to the European Benchmarks for the Chinese Language (EBCL). The mapping of TBCL to TOCFL shows that TBCL Levels 1 and 2 are below TOCFL entry level (A1) in listening, reading, and writing, and Level 1 is below TOCFL entry level (A1) in spoken language, indicating that TBCL has indeed achieved the goal of building a proficiency benchmark with Chinese language characteristics that is highly desired by Chinese language experts and scholars. The Ministry of Education's "Test of Chinese as Foreign Language" add the TBCL level to transcripts in 2021.

## (二) 華語文語料庫與能力基準整合應用系統 COCT&TBCL

為了提供使用者更容易的應用，本院將華語文分級標準分別建置「漢字分級標準檢索系統」、「詞語分級標準檢索系統」、「語法點分級標準檢索系統」等線上查詢系統。而且，將研發標準體系使用的「國教院索引典系統」、「華英雙語索引典系統」、「華語中介語索引典系統」、「語義場關聯詞查詢系統」、「教材編輯輔助系統」、「作文錯別字自動批改系統」等系統開放使用。

In order to provide users with an easier application, the Academy has set up online query systems such as "Chinese character grading scales retrieval system", "Word grading scales retrieval system", and "Grammar grading scale retrieval system". In addition, the Academy will develop and make available to the public "National Academy for Educational Research Thesaurus System", "Chinese-English Bilingual Thesaurus System", "Chinese Interlanguage Thesaurus System", "Semantic Field Conjunctive Word Query System", "Teaching Materials Editing Assistance System", and "Automatic Correction System for Typos in Compositions", which are used in the standards system.

然而，各系統各有自己的網址，使用者不易應用，因此本院將其整合建置「華語文語料庫與能力基準整合應用系統」(網址：<https://coct.naer.edu.tw>)，提升使用者應用的效能。

However, each system has its own URL, so it is not easy for users to apply. Therefore, the Academy has integrated them and set up the “COCT & TBCL” (URL: <https://coct.naer.edu.tw>) to enhance the effectiveness of user application.

本平台除了單一入口服務，並具備下述多面向服務優點，提供華語文教學、教材設計、測驗評量及研究應用。

In addition to the single portal, the platform also offers the following multi-faceted services and advantages, providing Chinese language teaching, materials design, testing and evaluation, and research applications.

1. 具有華語文語料類別最完整、內容最新、數量最多的語料庫

1. It is the most complete and up-to-date Chinese language corpus with the largest amount of content

本平台建置有華語書面語、口語、中介語及華英雙語等四類語料庫，為臺灣華語文語料類別最完整語料庫；這四類語料庫建置的語料都是最新，且數量最多。

其中，書面語語料庫內容涵蓋 2008 年以後，哲學及宗教類、科學類、應用科學類、社會科學類、史地類、語言文學類、藝術類、商業及金融類、休閒類等多元文章。口語語料庫內容則涵蓋 2008 年以後之法政軍事、財經、時事、科學、生活時尚、文教藝術等多元電視節目的語音及文字。由於，語料最新且數量最多，可使詞語及語法點的使用頻率、可能義項及前後搭配詞的分析，更具信效度並符合華語文目前實際語用情境。

中介語語料庫內容包括臺灣各大學華語文中心的華語學習者提供的作文，及國家華語測驗推動工作委員會授權的學生考試語料，內容包括題目、性別、文體(口語、書面語)、作者母語、能力等級、性質(限時考試、不限時作業)等資訊。華英雙語語

料庫內容涵蓋臺灣特色的文學、科學、財經、藝術、思想、文化、全球、休閒等議題。

The platform is equipped with four types of corpora: written, spoken, interlanguage, and Chinese-English bilingual. It is the most complete corpus in the category of Chinese language in Taiwan; all four corpuses are up-to-date and the largest in number.

The written language corpus covers a wide range of articles in philosophy and religion, science, applied science, social science, history and geography, language and literature, art, business and finance, and leisure from 2008 onwards. The spoken language corpus covers the audio and text of a wide range of television programs from 2008 onwards in legal, political and military affairs, finance, current events, science, lifestyle and fashion, literature, education and art. etc. With the most up-to-date and abundant corpus, the frequency of use, possible meanings and collocations of words can be analyzed more reliably and consistently with the current actual context of Chinese language use.

The content of the interlanguage corpus includes compositions by Chinese language students at the Chinese language centers of universities in Taiwan and student test materials authorized by the Steering Committee for the Test Of Proficiency-Huayu, including topic, gender, style (spoken, written), author's native language, proficiency level, nature (limited time tests, unlimited time assignments) and other information. The Chinese-English bilingual corpus covers topics such as literature, science, finance and economics, art, idea, culture, global, and leisure with Taiwan characteristics.

2. 具有線上查詢漢字、詞語及語法分級標準

2. It has online query for the grading scales of Chinese characters, words and grammar

華語文分級標準包括三等七級漢字 3,100 個、詞語 14,425 個及類詞綴 73 個，也包括第 1 級至第 5 級 496 個語法點。數量龐大，不易使用。因此，本平台建置「漢字分級標準檢索系統」、「詞



語分級標準檢索系統」及「語法點分級標準檢索系統」，提供使用者查詢漢字、詞語及語法點的等級，漢字、詞語在書面語及口語中的使用頻率，及語法點例句。

The Chinese language grading scales include 3,100 Chinese characters, 14,425 words, and 73 affixes in this 3 Grades and 7 Levels, as well as 496 grammar from Levels 1 to 5. The amount is huge and not easy to use. Therefore, the platform has established the “Chinese character grading scales retrieval system”, the “Word grading scales retrieval system”, and the “Grammar grading scales retrieval system”, which enable users to query the level of Chinese characters, words and grammar, the frequency of use of words in written and spoken language, and example sentences of grammar.

### 3. 提供完整技術報告

### 3. It provides complete technical reports

本院將本計畫歷年研發成果，出版《遣辭用「據」——臺灣華語文能力第一套標準》、《臺灣華語文語料庫——華語文教與學的必備工具》、《說情話境——華語文詞語情境分類》、《解詞造句——華語文基礎詞語彙編》、《跨域趨勢——臺灣華語文語料庫與能力基準整合應用系統》等 5 本技術報告，並建置於本平台，提供下載。

The Academy has published five technical reports, including *Weigh One's Words with Classical Allusion : The First Set of Standards for Chinese Language Ability in Taiwan*, *Corpus of Contemporary Taiwanese Mandarin—Tools for Teaching and Learning Mandarin Chinese*, *Conversational Scenarios—A Situational Classification of Mandarin Chinese words*, *Analyzing Words, Creating Sentences—Basic Mandarin Chinese Vocabulary*, *Interdisciplinary Trends — Integrated Applications of Corpus of Contemporary Taiwanese Mandarin And Taiwan Benchmarks for the Chinese Language* on the research and development results of this project over the years, and has disclosed them on the platform for download.

4. 整合應用語料庫、能力基準及自然語言技術並單一入口服務

4. It integrates COCT, TBCL and natural language technology with a single portal

除了上述語料庫及能力基準的應用系統，本院更將上述內容整合應用自然語言技術，建置「語義場關聯詞查詢系統」、「教材編輯輔助系統」、及「作文錯別字自動批改系統」等系統。不過，各系統各自有自己的網址，使用者不易記住，因此本院將這些系統整合建置單一入口網站（網址：<https://coct.naer.edu.tw>），提供使用者一站式的服務。包括：

- a. 查詢華語文漢字、詞語及語法點的等級，可分別使用「漢字分級標準檢索系統」、「詞語分級標準檢索系統」、「語法點分級標準檢索系統」。
- b. 分析詞語及語法點的使用頻率、可能義項及前後搭配詞，可應用「國教院索引典系統」。
- c. 分析華語非母語人士使用詞語、語法點及前後搭配詞的可能偏誤，可應用「華語中介語索引典系統」。
- d. 分析詞語的近義詞及反義詞，可應用「語義場關聯詞查詢系統」。
- e. 分析例句或文章所用詞語的等級，可應用「教材編輯輔助系統」。
- f. 分析作文錯別字並建議正確用字，可應用「作文錯別字自動批改系統」。
- g. 分析華英雙語的雙向翻譯、使用頻率、可能義項及華英雙語前後搭配詞，可應用「華英雙語索引典系統」。

In addition to the aforementioned COCT and TBCL benchmarks application systems, the Academy has integrated and applied the aforementioned content and natural language technology have built a “Semantic Field Conjunctive Word Query System”, “Teaching Materials Editing Assistance System”, and “Automatic Correction System for Typos in Compositions” and other systems. However, each system has its own URL, so it is not easy for users to remember. Therefore, the Academy has integrated them and set up the “Corpus of Contemporary Taiwanese

Mandarin (COCT)” (URL: <https://coct.naer.edu.tw>) to enhance the effectiveness of user application. Including:

- a. “Chinese character grading scales retrieval system”, “Word grading scales retrieval system”, and “Grammar grading scale retrieval system” can be used to query the level of Chinese characters, words and grammar.
- b. The frequency of use, possible meanings, and collocations of words can be analyzed using “National Academy for Educational Research Thesaurus System”.
- c. To possible errors in the use of words, grammar and collocations of words by non-native speakers of Chinese can be analyzed using “Chinese Interlanguage Thesaurus System”.
- d. Synonyms and antonyms of words can be analyzed using “Semantic Field Conjunctive Word Query System”.
- e. The level of words used in example sentences or articles can be analyzed using “Teaching Materials Editing Assistance System”.
- f. Typos can be analyzed and the correct use of words can be suggested using “Automatic Correction System for Typos in Compositions”.
- g. The two-way translation, frequency of use, possible meanings, and Chinese-English bilingual collocations of words can be analyzed using “Chinese-English Bilingual Thesaurus System”.

本系統使用點擊數，於 2021 年 1 月至 2021 年 12 月，已達 562 萬人次，大約去年同期的兩倍。

From January 2021 to December 2021, the hit rate of the system reached 5.62 million, almost twice the number for the same period last year.