

# ***Speaking Skills***

***in the Academic Context***



## **Contributors:**

Dr. Chen Hsueh Chu Rebecca

Ms. Chan Ching Hang Justine

Ms. Han Qianwen Veronica

Ms. Tian Jingxuan

Ms. Zhou Xiaona Tina



## Foreword

The purposes of this book are to introduce an English learners' corpus – The English Speech Corpus with Different Proficiency Levels and guide learners how to use the corpus to improve spoken English.

Since there are a number of established and recognized English language tests worldwide, such as TOEFL, IELTS and Cambridge English Qualifications, after reviewing the marking scheme of the tests, this book has taken IELTS as the example to demonstrate skills and strategies for handling the academic speaking test.

Chapter 1 will introduce the test format and marking criteria of the academic speaking test, as well as the corpus and learning website developed for learners and teachers on improving English speaking proficiency. Chapter 2 to Chapter 5 will be devoted on the explanations of each marking criteria, namely **Fluency and Coherence (FC)**, **Lexical Resource (LR)**, **Grammatical Range and Accuracy (GA)**, and **Pronunciation (PN)** respectively. Positive features and ways to avoid common mistakes will be introduced. Corpus-based activities are available along with the features and mistakes to enhance learners' experience. At the end of each chapter, *Teaching Corner* is provided for those who would like to teach the skills in their future career.

## Chapter 1: Introduction to Academic Speaking test and the corpus

By the end of this chapter, you will be able to:

- Understand the format and marking criteria of the IELTS speaking test
- Improve your speaking with the resources on the learning website
- Use the ‘Browse’ and ‘Search’ functions of the corpus

In this chapter, we will start by introducing the format and marking criteria of the speaking part of an internationally recognized English Language Test, IELTS. Then, resources available on the learning website, such as a question bank and mini quizzes, will be presented. Finally, you will learn how to use the ‘Browse’ and ‘Search’ functions of the corpus.

### Part I. About the IELTS speaking test

(All contents below are retrieved from: <https://www.ielts.org/for-test-takers/test-format>)

#### A. Test format

The IELTS speaking test is an oral interview between the test-takers and an examiner, and usually lasts for **11–14 minutes**. There are three parts, and the test assesses candidates’ use of spoken English. Every test is recorded.

#### Part 1 – Introduction and Interview

Timing: 4 – 5 minutes

Focus: to test test-takers’ ability to **communicate opinions and information** on everyday topics and common experiences or situations

Task type: The examiner asks general questions about test-takers themselves and a range of familiar topics, such as home, family, work, studies and interests.

#### Part 2 – Long Turn

Timing: 3–4 minutes (including one-minute preparation time)

Focus: to test test-takers’ ability to **speak at length on a given topic** (without further prompts from the examiner), using appropriate language and organising ideas coherently

Task type: The examiner gives the test-takers a task card which asks the test-takers to talk about a particular topic, includes points to cover in their talk and instructs the test-takers to explain one aspect of the topic. Test-takers are given one minute to prepare their talk, and are given a pencil and paper to make notes. The examiner asks the test-takers to talk for 1 to 2 minutes, stops the test-takers after 2 minutes, and asks one or two questions on the same topic.

#### Part 3 – Discussion

Timing: 4 – 5 minutes

Focus: to test the ability to **express and justify opinions and to analyse, discuss and speculate** about issues

Task type: Test-takers will be asked further questions about the topic in Part 2. These will give test-takers the opportunity to discuss more abstract ideas and issues in greater depth.

### *B. Marking Criteria*

Test-takers' speaking performances are assessed by certificated IELTS examiners based on the [official performance descriptors](#) which describe spoken performance at the nine IELTS bands. Four sub-skills are examined, namely **Fluency and Coherence (FC)**, **Lexical Resource (LR)**, **Grammatical Range and Accuracy (GA)**, and **Pronunciation (PN)**.

#### Fluency and Coherence (FC)

Fluency and Coherence in IELTS speaking test refers to the ability to talk with normal levels of continuity, rate and effort and to link ideas and language together to form coherent, connected speech. The key indicators involved are appropriate speech rate, appropriate speech continuity, logical sequencing of sentences, clear marking of stages in a discussion, narration or argument, and the proper use of cohesive devices (e.g., connectors, pronouns and conjunctions) within and between sentences.

Key indicators used for determining speakers' performance in FC include Repetition, Self-correction, Discourse markers and Connectives, Speech length, Speech rate, and the number and duration of filled pause, silent pause and silent pause at the end of a phrase, a clause or a sentence.

#### Lexical Resource (LR)

In this part, the range of vocabulary used and the precision with which meanings and attitudes can be expressed are evaluated. Several key indicators are involved, such as the variety of words used, the adequacy and appropriacy of the words used and the ability to circumlocute (get round a vocabulary gap by using other words) with or without noticeable hesitation.

Key indicators used for determining speakers' performance in LR include less common words, paraphrase, idiomatic expressions as well as appropriate collocations and word choice.

#### Grammatical Range and Accuracy (GA)

Grammatical Range and Accuracy in IELTS speaking test refers to the range and the accurate and appropriate use of the test takers' grammatical resource. The key indicators involved are the length and complexity of the spoken sentences, the appropriate use of subordinate clauses, and the range of sentence structures, especially to move elements around for information focus, the number of grammatical errors in a given amount of speech, and the communicative effect of error.

Key indicators used for determining speakers' performance in GA include grammatical errors, and the use of sentence structures such as coordinate clause, adverbial clause, relative clause, noun clause and inversion.

## Pronunciation (PN)

In this part, the ability to produce comprehensible speech to fulfil the Speaking test requirements is examined. Several key indicators are involved, such as the amount of strain caused to the listener, the amount of the speech which is unintelligible, and the noticeability of L1 influence.

Key indicators used for determining speakers' performance in PN include consonants and vowels features, mispronunciation, the use of consonant-vowel linking, as well as appropriate intonation, word stress and rhythm.

### *What is a 'corpus'? How can it help my language learning?*

A corpus is "a large, principled collection of naturally occurring texts (written or spoken) stored electronically" (Reppen, 2010, p.2). There are many types of corpora compiled worldwide. Some of them contain language data from native speakers, such as [British National Corpus \(BNC\)](#), while some of them contain data from learners, such as The English Speech Corpus with Different Proficiency Levels introduced in this book.

By analysing the authentic data produced by other language learners, you can identify the common errors and positive features more easily and thus be aware of the mistakes that you may also have produced throughout the learning.

## **Part II. The corpus**

### *A. About the corpus*

[The English Speech Corpus with Different Proficiency Levels](#) aims to provide learners, teachers and researchers with high-quality authentic recordings from practice tests and detailed annotations of linguistic features of learners with different English proficiency levels. The corpus contains 72 sets of data from free speeches. Of the 72 sets of data, 42 are collected from Chinese Mainland and Hong Kong learners, and 30 are retrieved from IELTS speaking official videos.

There is currently over 300 minutes of recording data with linguistic annotations that focus on four aspects (i.e., Fluency and Coherence, Lexical Resource, Grammatical Range and Accuracy, and Pronunciation) according to the IELTS speaking criteria, which can help Chinese English learners identify difficulties in English speaking learning. The high-quality recordings are ideally suited for learners, teachers, and researchers in the Greater China.

To find a recording of a practice test or a video of the IELTS speaking test sample, you can use either the "Browse" or "Search" function. This site is fully searchable, not just by browsing speakers' biographic information, like the gender, the age, the IELTS Speaking score they achieved and the IELTS speaking session they took; you can also use the "Search" function to search the recording samples by selecting the linguistic features you are interested in.

All the annotations that have been added for the current recordings are solely our suggestions. We do not claim that these are 100% accurate or the only way the speech can be analysed. All raw recordings and transcripts are open for download.

EdUHK Home Our English Speech Corpus Classroom Presentation For Learners For Teachers Learning Resources

## Our English Speech Corpus

**OVERVIEW**

The English Speech Corpus with Different Proficiency Levels is expanded and redeveloped from the previous small-scale spoken corpus. It contains 72 sets of spontaneous speech data and 15 sets of classroom presentation data. Of the 72 sets of data, 42 are collected from mainland China and Hong Kong learners, and 30 are retrieved from IELTS speaking official videos.

In addition to the authentic speech data, detailed annotations were also made based on the four aspects of IELTS speaking criteria (i.e., fluency and coherence, grammatical range and accuracy, and pronunciation), which can help identify difficulties in English speaking learning for Chinese English learners with different language levels. The high-quality recordings are ideally suited for teachers, learners, and researchers around the world.

### B. How to use the 'Browse' function?

Having entered the corpus page, you will be directed to the 'Browse' page where you can use the 'Browse' function.

香港教育大學 The Education University of Hong Kong Home How To About Browse Search Resources

## Browse

### Browse Speakers

Gender : [Either] Age : [ANY] IELTS Speaking Score : [ANY] Speaking Session : [ANY] [Filter Speaker](#)

Remarks:  
 Each speaker has four data pages in each speaking session:  
 FC for features of *fluency and coherence*,  
 LR for features of *lexical resources*,  
 GA for features of *grammatical range and accuracy*,  
 PN for features of *pronunciation*.

IELTS Speaking Score: 5.0												
Part 1	1				2				3			
	FC	LR	GA	PN	FC	LR	GA	PN	FC	LR	GA	PN
Part 2	4				5				6			
	FC	LR	GA	PN	FC	LR	GA	PN	FC	LR	GA	PN
	7											
Part 3	FC	LR	GA	PN								
	8				9				10			
	FC	LR	GA	PN	FC	LR	GA	PN	FC	LR	GA	PN
	11											

© The Education University of Hong Kong (EdUHK) March 2021

There are two ways for you to get into the data page where we put the transcripts and annotations of the speakers' performance.

## 1. Direct browsing

The 'Browse' page has shown all data pages of each speaker. They are categorized based on the speaking score (i.e., from 5.0 to 9.0) and speaking session (i.e., Part 1, Part 2 and Part 3). Each speaker contains 4 data pages, namely FC, LR, GA and PN. These 4 pages are in line with the marking criteria of the speaking test.

You can simply click on the FC/LR/GA/PN page under each speaker according to the band score and session you want to look at. A plain transcript will be shown after the click. Below is an example from the FC page of speaker 1, who scored 5.0 in the test. On the left of the transcript, you can find an info box containing details about the speaker (e.g., gender, IELTS Speaking Score) and the transcript (i.e., IELTS Speaking Session). On the transcript, audio segments are available after each question and response.

The screenshot displays the 'Fluency & Coherence' page for Speaker 1. On the left, an 'Info box' provides details: Speaker: 1, Gender: Female, IELTS Speaking Score: 5, and IELTS Speaking Session: Part 1. The main area shows a transcript with audio player controls for each segment. A callout box labeled 'Audio files of the question / response' points to the audio player for the first response. The transcript includes the following text:

E-examiner S-student  
E: OK, so, first of all, I am going to ask some questions about yourself. Let's talk about what you do. Do you work or are you a student? 0:00 / 0:08  
S: I'm a, I'm, I'm an undergraduate student studying in XX university and major in Business English. 0:00 / 0:11  
E: Why did you choose to study this? 0:00 / 0:03  
S: Because in, in my high school, I see many English movies and I ... I'm so very big fan of English. So, I want to ju ... I want do something about this. So, I choose English. And also before contents, b , business is very easy to find a job. I... So, I chose it. 0:00 / 0:30  
E: Do you prefer to study in the morning or in the afternoon? 0:00 / 0:03  
S: I prefer to study in the morning because in the morning, I... it the beginning of the day and I

If you want to read the annotations, simply scroll down to look for the feature list:

Select the feature(s) that you want to learn, the annotations will be automatically displayed in the transcript. You can also select **All** to enable all annotations, or **None** to disable all annotations. When you browse the data page of speakers with high band scores (i.e., 8.0 or above), at the bottom of the transcript, there is a link to another page showing a combined version of annotation. The annotations on Fluency and Coherence and Pronunciation will be combined, while annotations of Lexical Resource and Grammatical Range and Accuracy will be combined.

Note:  
Click the icon  to view FC&PN-in-one version of annotation.

**Feature List**

Pronunciation

- Dental fricatives /θ/ & /ð/
- Aspiration & unaspiration
- Omission of consonant(s)
- Insertion of vowel(s)
- Omission of vowel(s)
- Mispronunciation
- Word stress
- Consonant-vowel linking

All / None

Check the combined annotation

Select the feature(s) you want to learn

At the end of the data page, there is also a PDF file summarizing the annotated features. The PDF file also provides general comments and some suggested corrections. Details provided for each marking criteria are as follow:

Marking criteria	Details
Fluency and Coherence	<ul style="list-style-type: none"> <li>• The occurrence of repetition, self-correction, discourse markers &amp; connectives, pausing</li> <li>• Speech length, speech rate</li> <li>• Duration and occurrence of filled pausing, silent pausing, and silent pausing – E</li> </ul>
Lexical Resource	<ul style="list-style-type: none"> <li>• The occurrence of appropriate less common words, idiomatic expressions, paraphrase and collocations</li> <li>• Suggested corrections for inappropriate collocations and word choices</li> <li>• Distribution of words according to BNC-COCA type list</li> </ul>
Grammatical Range and Accuracy	<ul style="list-style-type: none"> <li>• General comments on the use of sentence structures and grammatical errors</li> <li>• Suggested corrections for grammatical errors</li> </ul>

## Pronunciation

- The occurrence of segmental features, including consonants and vowels
- General comments on suprasegmental features, including intonation, stress, linking and rhythm

You can also download this file for future reference.

The screenshot shows a document viewer with the following content:

**General Description**  
Consonant features: in purple  
Vowel features: in green

Intonation: Intonation is generally appropriate. Both rising intonation and falling intonation are used.

Stress: Primary stresses are put on appropriate syllables, except one word:  
ANCESTORS → anCEStors.

Linking: Linking is occasionally used (highlighted in light green).

Rhythm: The duration of unstressed function words is shorter than the stressed syllables in content words. However, the tonic syllables have similar duration with stressed syllables in content words.

e.g.  
Student:  
... that is the most important because sometimes you might encounter some culture  
O o O o O o O o O o o o O o O o  
shock ...  
O

Suggested:  
... that is the most important because sometimes you might encounter some culture  
O o O o O o O o o o o o O o O o  
shock ...  
O

© The Education University of Hong Kong (EdUHK) | March 2021

## 2. Browse with filter

Instead of manually scrolling the 'Browse' page, you can look for the desired pages by applying filters such as gender, age, IELTS speaking score and speaking session.

香港教育大學  
The Education University  
of Hong Kong

Home How To About Browse Search Resources

/ Browse

### Browse Speakers

Gender :  Age :  IELTS Speaking Score :  Speaking Session :

**Remarks:**  
Each speaker has four data pages in each speaking session:  
FC for features of *fluency and coherence*,  
LR for features of *lexical resources*,  
GA for features of *grammatical range and accuracy*,  
PN for features of *pronunciation*.

IELTS Speaking Score: 7.0												
Part 3	56				57							
	FC	LR	GA	PN	FC	LR	GA	PN	FC	LR	GA	PN
	59											
	FC	LR	GA	PN								

Click [here](#) to browse raw data

© The Education University of Hong Kong (EdUHK) | March 2021

### C. How to use the 'Search' function?

Apart from the 'Browse' page, the corpus also comes with a 'Search' function.

香港教育大學  
The Education University  
of Hong Kong

Home How To About Browse Search Resources

Search

### Search Features

Band :

Category :

Step 1 :

Step 2 :

© The Education University of Hong Kong (EdUHK) | March 2021

This search engine is very straight-forward. All you need to do is select the **Band** (i.e., from 5.0 to 9.0) and **Category** (i.e., the four marking criteria). Then, the engine will show the feature list of the selected category in **Step 1** for you to choose. You can also select NONE SELECTION as the choice to show results containing all features.

If the feature that you have chosen contains sub-category, you can further refine the search result in **Step 2**.

**Search Features**

Band : 5.5

Category : Fluency & Coherence

Step 1 : Silent Pause (SP)

Step 2 : NONE SELECTION

Search Result

IELTS Speaking Score: 5.5

		12				13				14			
Part 1		FC				FC				FC			
		15				16							
		FC				FC							
Part 2		19				20				21			
		FC				FC				FC			
		22											

All search results will be displayed with the same presentation style as shown in 'Browse' page. The only difference is that when you enter a particular data page, the feature that you have chosen in the 'Search' page will be displayed on the transcript. Therefore, you don't need to scroll down and select the feature from the feature list.

The screenshot shows the website interface for the English Speech Corpus. At the top, there is a navigation menu with links: Home, How To, About, Browse, Search, Resources, and Login. The logo of The Education University of Hong Kong is in the top left. The main content area is titled 'Fluency & Coherence'. On the left, there is a box with speaker information: Speaker: 12, Gender: Female, IELTS Speaking Score: 5.5, IELTS Speaking Session: Part 1. The main area contains a dialogue between an examiner (E) and a student (S). The dialogue includes several lines of speech with playback controls. A yellow callout box highlights a specific line of student speech: 'S: Because I love this major and s, since I am (SP) child, I (SP), I love to communicat, (0.10s) (0.29s) communication with my friend and my family. So, I choose this major.' A speech bubble labeled 'Annotated transcripts' points to this highlighted text. At the bottom of the page, there is a copyright notice: © The Education University of Hong Kong (EdUHK) | March 2021.

#### D. How to cite the corpus?

If you want to cite this corpus in your assignments / dissertations / teaching materials, please use the following information: (citation style: APA 7<sup>th</sup> edition)

Chen, H.C. (2022). *The English Speech Corpus with different proficiency levels*. The Education University of Hong Kong. Retrieved from <https://lml-learning.eduhk.hk/englishspeechcorpus/index.aspx>.

### Part III. Learning website

In the learning website (accessible [here](#)), you can find the access to the corpus, recordings of classroom presentation and various resources for learners and teachers.

EdUHK Home Our English Speech Corpus Classroom Presentation For Learners For Teachers Learning Resources

## Corpus-based English Speaking Teaching and Learning

Improve your English Speaking Skills @ English Speech Corpus

### About the project

"Development of a flipped corpus-aided spoken English training platform for Chinese learners of English"

In this project, we developed a corpus-aided spoken teaching and learning platform which aims to help teachers, learners and researchers have a better understanding of the spontaneous speech features and problems that Chinese English learners have. An English Speech corpus and an online flipped corpus-aided spoken English training programme are the key elements of this project. Over 80 sets of spontaneous data, including IELTS spoken speech and classroom presentation data produced by learners with different English proficiency levels, were collected, transcribed and annotated for four features (fluency and coherence, lexical resource, grammatical range and accuracy, and pronunciation). Learners and language teachers can identify the potential spoken features they or their students may have through getting access to the authentic data as examples for their spoken English learning.

In **Classroom Presentation** page, you will find authentic recordings from students of different English proficiency levels (i.e., IELTS speaking 5, 5.5, 6, 6.5, 7, and 7.5) who have delivered speeches on topics related to **Life Education**, such as *A time you feel satisfied with Life*, *A time you dealt with stress*, etc. Evaluation forms are provided in the Activity session of the **For Learners** page.

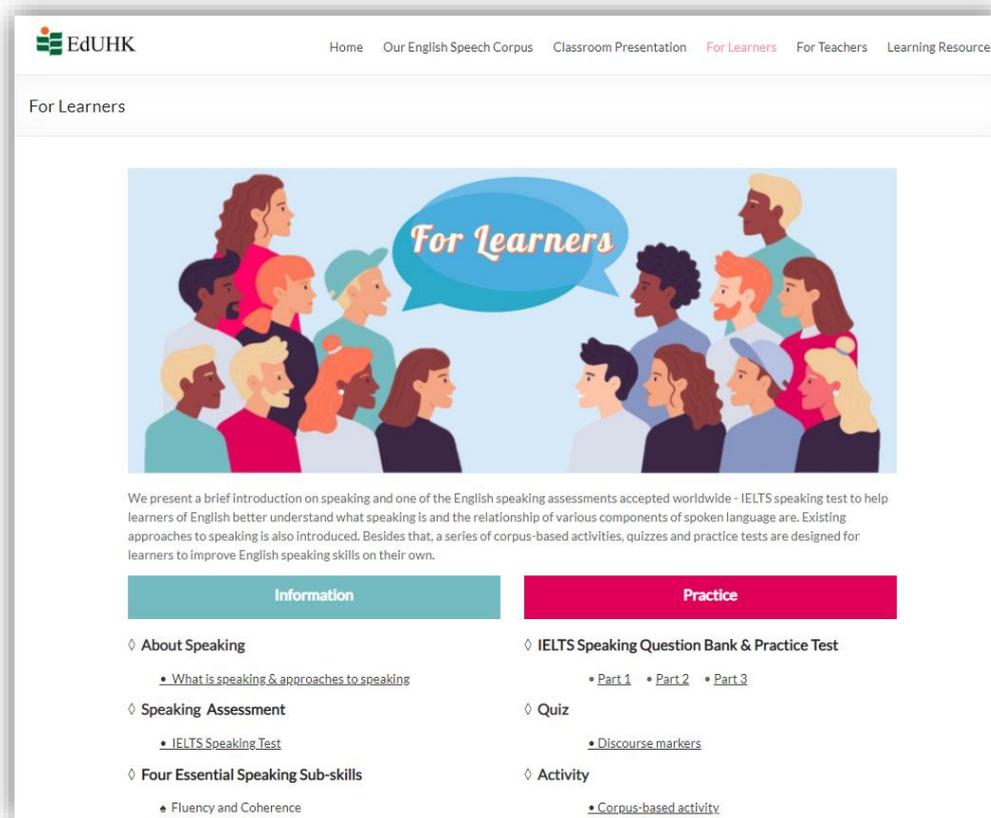
EdUHK Home Our English Speech Corpus Classroom Presentation For Learners For Teachers Learning Resources

## Classroom Presentation

There are 15 5-minute classroom presentations, which were presented by 15 mainland Chinese learners with different English proficiency levels. The topics of the presentation are mainly on life education (e.g., a time you feel satisfied with life, something you did to improve your mindfulness, methods of emotion regulation that you learned from others, etc). Each learner elaborated the topic based on their own experience, such as songs they listened, movies they watched, and books or articles they read.

IELTS Speaking Score	Classroom Presentation			
5.0	1	2	3	
5.5	4	5	6	7
6.0	8	9		

In the **For Learners** page, there are information on the four marking criteria of the speaking test, as well as different forms of practice including question bank and quizzes. **Corpus-based activities** and evaluation forms of **classroom presentation** are available. For the classroom presentation session, after listening to the presentations in the **Classroom Presentation** page, you can evaluate their performance with forms and indicators provided.



EdUHK Home Our English Speech Corpus Classroom Presentation **For Learners** For Teachers Learning Resources

For Learners



We present a brief introduction on speaking and one of the English speaking assessments accepted worldwide - IELTS speaking test to help learners of English better understand what speaking is and the relationship of various components of spoken language are. Existing approaches to speaking is also introduced. Besides that, a series of corpus-based activities, quizzes and practice tests are designed for learners to improve English speaking skills on their own.

Information	Practice
<ul style="list-style-type: none"><li>◇ About Speaking<ul style="list-style-type: none"><li>• <a href="#">What is speaking &amp; approaches to speaking</a></li></ul></li><li>◇ Speaking Assessment<ul style="list-style-type: none"><li>• <a href="#">IELTS Speaking Test</a></li></ul></li><li>◇ Four Essential Speaking Sub-skills<ul style="list-style-type: none"><li>• <a href="#">Fluency and Coherence</a></li></ul></li></ul>	<ul style="list-style-type: none"><li>◇ IELTS Speaking Question Bank &amp; Practice Test<ul style="list-style-type: none"><li>• <a href="#">Part 1</a> • <a href="#">Part 2</a> • <a href="#">Part 3</a></li></ul></li><li>◇ Quiz<ul style="list-style-type: none"><li>• <a href="#">Discourse markers</a></li></ul></li><li>◇ Activity<ul style="list-style-type: none"><li>• <a href="#">Corpus-based activity</a></li></ul></li></ul>

In the **For Teachers** page, you can find the Corpus-aided Teaching Training Programme with materials and video clips open for free access.

EdUHK Home Our English Speech Corpus Classroom Presentation For Learners **For Teachers** Learning Resources

For Teachers



**Common linguistic features of Chinese learners with different English proficiency levels**

We summarized salient linguistic features of learners from Hong Kong and mainland China based on data from the corpus. The features from four aspects (fluency and coherence, lexical resource, grammatical range and accuracy, and pronunciation) are quite different among speakers with different IELTS speaking band scores. English teachers are strongly encouraged to know their students' difficulties in English speaking, set their focus on teaching speaking and design teaching materials based on our summary.

- » [Fluency and Coherence](#)
- » [Lexical Resource](#)
- » [Grammatical Range and Accuracy](#)
- » [Pronunciation](#)

**Corpus-aided Teacher Training Programme**

The programme helped pre- and in-service teachers learn about an innovative corpus-aided approach. All the training materials and recordings of the programme are accessible.

- » [Training Materials](#)
- » [Video Clips](#)

Finally, references and other online resources are listed in the **Learning Resources** page. Useful websites are identified and categorised based on the four marking criteria for your further self-learning.

Learning Resources



Fluency and Coherence

• Discourse marker list

<https://dictionary.cambridge.org/grammar/british-grammar/discourse-markers-so-right-okay>

Lexis

• Online Collocation Dictionary

<https://www.freecollocation.com/>

<https://ozdic.com/>

<https://www.macmillandictionary.com/browse/collocations/british/a/>

• Online Idiom Dictionary

<https://www.theidioms.com/>

Grammar

Pronunciation

<https://corpus.eduhk.hk>



**IMPROVE SPEAKING SKILLS**  
**@ English Speech Corpus**



[https://corpus.eduhk.hk/english\\_speech\\_corpus/](https://corpus.eduhk.hk/english_speech_corpus/)

