PISA

PISA 2025 Foreign Language Assessment Framework

OECD

PISA 2025 Foreign Language Assessment Framework

Please cite this publication as:
OECD (2021), *PISA 2025 Foreign Language Assessment Framework*, PISA, OECD Publishing, Paris.

This document was produced with the financial assistance of the European Union. The views expressed herein can in no way be taken to reflect the official opinion of the European Union.

Table of contents

Foreword

The OECD Programme for International Student Assessment (PISA) assesses the extent to which 15-year-old students near the end of their compulsory education have acquired the knowledge and skills essential for full participation in modern societies. The assessment does not just ascertain whether students can reproduce knowledge; it also examines how well students can extrapolate from what they have learned and can apply that knowledge in unfamiliar settings, both in and outside of school. This approach reflects the fact that modern economies reward individuals not for what they know, but for what they can do with what they know. The triennial assessment, launched in 1997, focuses on the core school subjects of reading, mathematics and science. Optional components are also offered.

The PISA 2025 cycle will include, for the first time, an optional assessment of foreign languages, which will be offered every two PISA cycles, to allow for the analyses of trends. The assessment will provide policy makers and educators with comparable results of their students' foreign language competence and allow them to gain insights into the best practices and policies for teaching and learning a foreign language. The first cycle will start with the assessment of English and will focus on three skills: reading, speaking and listening. The coverage may progressively widen in future cycles.

This publication presents the guiding principles behind the PISA 2025 Foreign Language Assessment. It discusses the relevance of assessing foreign languages and provides the definition of foreign language proficiency that is used to guide the assessment. The publication also presents the theory underlying the tests to assess foreign language proficiency by describing the cognitive processes that learners employ when using a foreign language according to their proficiency level. It also discusses how the skills of reading, listening and speaking are assessed. Further, it outlines the framework for the various questionnaires distributed to students, school principals, parents and teachers, and the framework for the Foreign Language Assessment system-level questionnaire. It concludes with the description of how the assessment results will be reported.

Acknowledgements

The PISA Foreign Language Assessment framework is the product of a collaborative effort between the countries and economies participating in PISA, national and international experts and institutions and the OECD Secretariat. It was drafted by the OECD Secretariat with the guidance and input of expert groups.

The European Commission co-funded the framework and assisted with expertise built on previous work at the EU level, and integrating the Council of Europe´s Common European Framework of Reference for language competences. Kristina Cunningham provided input and co-ordinated technical feedback from the European Commission and the Council of Europe.

Chapter 3 was developed in collaboration with Cambridge Assessment English. The sections of the framework related to the cognitive framework, the proficiency levels and the test items that will be provided by Cambridge Assessment English for use in the PISA 2025 Foreign Language Assessment were developed with contributions from Cambridge Assessment English.

The PISA Governing Board steered the development of the framework. It was overseen by Andreas Schleicher and Yuri Belfali, guided by Tue Halgreen and managed by Catalina Covacevich.

Chapter 1 was drafted by Catalina Covacevich, Carla Campos and Gabriele Marconi, and edited by Marilyn Achiron. This chapter was co-funded by the European Commission.

Chapter 2 was drafted by Catalina Covacevich and Jimena Vargas, with contributions from Tue Halgreen, Francesco Avvisati and Gabriele Marconi, and edited by Julie Harris. This chapter was co-funded by the European Commission.

Chapter 3 was drafted by Angeliki Salamoura (from Cambridge Assessment English), with contributions from Hanan Khalifa (from Cambridge Assessment English), Catalina Covacevich, Francesco Avvisati and Gabriele Marconi, and edited by Julie Harris.

Chapters 4 and 6 were drafted by Francesco Avvisati and Catalina Covacevich, with contributions from Gabriele Marconi, and edited by Julie Harris. These chapters were co-funded by the European Commission.

Chapter 5 was drafted by Gabriele Marconi, Carla Campos Cascales, Catalina Covacevich and Tue Halgreen, with contributions from Zsuzsa Bakk, Alfonso Echazarra, Pablo Fraser and Daniel Salinas, and edited by Marilyn Achiron. This chapter was co-funded by the European Commission.

Lisa-Maria Müller was the main expert advising on Chapters 1 and 2 and participated in the drafting of sections of these chapters. Additional input was provided by Talia Isaacs, Gad Lim, David Little, Lorena Meckes, José Noijons and Jean Francois Rouet.

The expert group for Chapters 3, 4, and 6 was formed by Talia Isaacs, Gad Lim, David Little, Lorena Meckes, José Noijons and Jean Francois Rouet.

The expert group for Chapter 5 was formed by Kathleen Graves, Haedong Kim, Gladys Quevedo-Camargo, Angeliki Salamoura and Sanneke Schouwstra, with additional input from Suzana Bitenc Peharc, Letizia Cinganotto, Steven Culpepper, Alister Cumming, Katrijn Denies, Janet Enever,

Elsa Fernanda González, Kulaporn Hiranbuara, Gisella Langé, Icy Lee, Patsy Lightbown, David Marsh and Lisa-Maria Müller. A list of the contributing experts and their institutional affiliations is found in Annex A.

Hanan Khalifa, Andrew Kitney, Sian Morgan, Emma Pathare, Darren Perrett, Martin Robinson and Angeliki Salamoura (from Cambridge Assessment English) contributed throughout the framework to content related to the cognitive framework, the proficiency levels and the test items.

Jimena Vargas co-ordinated the production of this publication, and Alison Burke provided communications assistance. Thomas Marwood and Hanna Varkki provided administrative support.

1 Policy interest in assessing foreign language learning

This chapter discusses how being able to communicate in more than one language has become a critical skill in today's globalised world, and describes the benefits that come with foreign language learning. It also describes countries' efforts to increase foreign language teaching. In addition to presenting a historical overview of various approaches to foreign language teaching around the world, this chapter underlines the policy relevance of assessing foreign language skills today, particularly in view of the increasing role of foreign language teaching and learning in most education systems.

Why learn a foreign language?

Globalisation, technological innovation and human migration flows have made interactions amongst people from different countries and cultures almost inevitable. In 2017, there were 260 million people living outside their country of birth, 1.3 billion international tourist arrivals and around 5 million students enrolled in a tertiary education programme outside their home country. All of these numbers have been steadily increasing in recent years (OECD, 2018[1]; OECD, 2018[2]; UNWTO, 2019[3]). International trade in goods and services accounts for well over half of global GDP (World Bank, 2019[4]). Given these developments, a large number of people need to reach some level of proficiency in more than one language to be able communicate and interact with others.

Education systems in many different countries and economies are striving to respond to this challenge. As illustrated in Box 1.1, over the past few decades more and more countries have emphasised the importance of teaching and learning foreign languages. While teaching foreign languages in general is increasingly central in education policies, English has become the main foreign language being taught and learned.

Box 1.1. Some examples of governments' efforts to increase foreign language teaching

- 1998: Korea introduces English at the primary level
- 2003: Chile initiates "English Opens Doors", a programme to strengthen foreign language learning
- 2008: Colombia introduces the Bilingualism National Plan (*Colombia Bilingüe*)
- 2010: The Abu Dhabi Education Council (ADEC) launch the "New School Model", implementing a bilingual English-Arabic model for curriculum in public schools
- 2011: Sweden establishes English as one of three core subjects in school, alongside mathematics and literacy
- 2017: Ireland sets up "Languages connect" to promote teaching of foreign language skills

As illustrated in Table 1.1, countries and economies are aware of the multiple benefits of learning more than one language and describe them in the official curriculum.

Table 1.1. Reasons for foreign language learning in primary and/or secondary education, selected education systems

Education system	Reasons for foreign language learning (as stated in the official curriculum)
The Flemish Community of Belgium	"The importance of knowing multiple languages increases as society becomes more multicultural and as countries are confronted with the effects of globalisation. A country that succeeds in bringing its population to a minimum level of multilingualism avoids the danger that monolingualism becomes a risk factor of social exclusion for the individual and of isolation in the world community" (translated from (Flemish Ministry for Education and Training, 2000[5]))
Costa Rica	"Learners need an updated curriculum that reflects the knowledge, skills, and abilities needed to communicate in a range of language use contexts and to succeed in the information age as 21st century learners. […] This view is in line with the concept of education for a new citizenship that maintains that 21st century learners must integrate proactively in a globalised world while strengthening their national and global identity" (República de Costa Rica, 2016[6])

Japan	"With globalisation advancing rapidly, it is assumed that communication skills in foreign languages are required not only in some industries and occupations as in the past, but also in various contexts and situations in daily life" (translated from (MEXT, 2017[7]))
New Zealand	"Learning a new language extends students' linguistic and cultural understanding and their ability to interact appropriately with other speakers. Interaction in a new language, whether face to face or technologically facilitated, introduces them to new ways of thinking about, questioning, and interpreting the world and their place in it. Through such interaction, students acquire knowledge, skills, and attitudes that equip them for living in a world of diverse peoples, languages, and cultures" (Ministry of Education - New Zealand, 2007[8])
United Kingdom	"Learning a foreign language is a liberation from insularity and provides an opening to other cultures. A high-quality language education should foster pupils' curiosity and deepen their understanding of the world. It should also provide opportunities for them to communicate for practical purposes, learn new ways of thinking and read great literature in the original language. Language teaching should provide the foundation for learning further languages, equipping pupils to study and work in other countries" (UK Department for Education, 2014[9])

The benefits of foreign language learning can be divided into three categories: intercultural understanding; economic benefits; and cognitive benefits. These are further described below.

Intercultural understanding

Language learning does not only improve communication but, is also an avenue towards understanding the diversity of human culture and languages (Fischer, 2012[10]). Through language, people can learn about and gain access to other cultures; enhance their cultural awareness and the understanding of other groups' values; and develop an awareness of cultural complexity (Curtain and Dahlberg, 2004[11]; Gudykunst, 2003[12]; Marian and Shook, 2012[13]).

Students of foreign languages learn that not everything is translatable, and can begin to appreciate how differently the world is perceived and understood by people from different cultures. Conversely, learning a foreign language can enable a person to discover his or her own cultural and personal identity through other cultural perspectives (Della Chiesa, Scott and Hinton, 2012[14]). Through its focus on communication with "others" who live beyond their linguistic communities, foreign language learning fosters intercultural communicative competencies and contributes to citizenship education (Porto, Houghton and Byram, 2018[15]).

The understanding of other cultures and visions of the world that is fostered by foreign language learning is an important prerequisite for active participation in a globalised world. The Council of Europe includes linguistic, communicative and plurilingual skills amongst the eight skillsets that are important in democratic cultures (Council of Europe, 2020[16]).

Economic benefits

Individuals who know one or more foreign languages, regardless of their proficiency, are more likely to be employed than monolingual speakers (Araújo et al., 2015[17]). Foreign language-proficient professionals also benefit from better career opportunities and labour mobility, and have greater chances of being promoted to higher-level jobs (Canadian Heritage, 2016[18]; European Commission, 2012[19]; Isphording, 2015[20]; Ahuja, Chucherd and Pootrakool, 2006[21]). There is also evidence that foreign language speakers earn higher salaries both in bilingual and non-bilingual countries, the returns varying according to the language and to the area of work (Garrouste, 2008[22]; Grin, 1999[23]). For example, a study from the United States shows that college graduates who speak a second language earn wages that are 2% higher, on average, than those of graduates who do not (Saiz and Zoido, 2002[24]), while in Morocco, the salary gap

between someone who speaks English and someone who doesn't can reach 12% (Euromonitor International, 2012[25]).

At the macro level, foreign language skills can contribute to a country's economy in a variety of ways. Countries with foreign language proficient populations trade more (Melitz, 2008[26]; Fidrmuc and Fidrmuc, 2016[27]). For example, it has been estimated that a lack of foreign language proficiency costs the United Kingdom GBP 48 billion per year, or 3.5% of its GDP (Foreman-Peck and Wang, 2014[28]). In addition, there is evidence that learning languages can be used as a dynamic tool to attract foreign direct investments (Kim et al., 2014[29]). Foreign language proficiency can also help dismantle language barriers across firms, making technical co-operation and innovation across borders easier (Canadian Heritage, 2016[18]; Tenzer, Terjesen and Harzing, 2017[30]).

Cognitive benefits

Some research suggests that students who learn a foreign language are more creative and better at solving complex problems than those who do not (Bamford and Mizokawa, 1991[31]). There is also evidence that studying a foreign language improves attention and mental alertness in adults after only one week of study (Woll and Wei, 2019[32]; Bak et al., 2016[33]). These abilities could be transferrable to other subjects.

Goethe wrote that "Those who do not know other languages know nothing of their own" (Della Chiesa, Scott and Hinton, 2012[14]). There is now scientific evidence to support this claim, as foreign language learning has been found to have a facilitative effect on first language literacy (Murphy et al., 2015[34]). Learning a foreign language also provides an opportunity to acquire metalinguistic skills and knowledge that can help in learning other languages (Rothman, Cabrelli and De Bot, 2013[35]).

Development of foreign language teaching

While governments have intensified their efforts to promote foreign language learning in recent decades, the teaching of foreign languages is not new. This section provides a brief overview of some foreign languages that have traditionally been taught, why they have been taught, and the teaching approaches used (McLelland, 2017[36]).

Lingue francae

A lingua franca is a common means of communicating for speakers of different first languages (Seidlhofer, 2011[37]; UNESCO, 1953[38]). Today English is considered a global lingua franca (Crystal, 1987[39]; Colombian National Ministry for Education, 2006[40]) and is the most used language on internet (Bokor, 2018[41]; Miniwatts Marketing Group, 2019[42]; Bada, 2018[43]). It is also currently the most widely taught foreign language (Crystal, 1987[39]; Ammon, 2015[44]).

Before English, other languages were used as linguae francae, albeit not as widely. Arabic, French, Hindi, German, Greek, Japanese, Latin, Malay, Mandarin Chinese, Nahuatl, Quechua, Russian, Spanish, Swahili and Tupi are among the languages that historically fulfilled the role of a lingua franca or continue to do so in some regions of the world. For example, modern Standard Arabic is used as a lingua franca in the Middle East and North Africa (European Commission, 2010[45]). It is also gaining in popularity as a foreign language to be studied and the British Council identified it as one of the top five Languages for the Future in a recent report (Tinsley and Board, 2017[46]).

While being a lingua franca plays an important role in the choice of the foreign languages to be studied, it is not the only reason and there is a widespread understanding that it is important to learn different foreign languages. In 2002, the European Union affirmed that one of the goals of quality education is to "improve

the mastery of basic skills, in particular by teaching at least two foreign languages from a very early age" (Council of the European Union, 2002[47]).

Overview of influential teaching approaches

Approaches to language teaching need to take two fundamental things into account: the desired outcome of language learning and the process of how to achieve this goal (Widdowson, 2004[48]). Historically, language teaching has alternated between different approaches, for example defining language as an object of analysis or a means to communicate (Celce-Murcia, 2014[49]).

While this section provides a broad overview of some of the most influential teaching approaches, this list is not exhaustive. It is also important to note that the history of language learning is not homogeneous. Approaches are likely to have co-existed, as they do today; and while the prevailing pedagogical orientations across countries and time are well documented, there is little insight into the practices adopted in the classroom (Musumeci, 2009[50]).

Early stages of formal foreign language teaching

Language learning has always been important as it allows people to trade and exchange with others. In Europe, accounts of formal language teaching typically start with the teaching and learning of Latin and Greek (Musumeci, 2009[50]). During the classical and medieval periods, first Greek then Latin, the two linguae francae at the time, were taught with the aim to turn the educated elite into fluent speakers, readers and writers of the languages (Celce-Murcia, 2014[49]).

In China, trade and diplomatic relations led to political attention to the teaching of foreign languages, and to the creation of a first foreign language school in 1289, where Persian was taught to young nobility. In the 14th century other schools were established, other languages as Manchu, Tibetan, Dai language, Uighur, Burmese and Thai started to be taught, and the first grammar book to learn how to read and write in Mongol was published. The learning of vocabulary and the practice of translation were the main methods to learn foreign languages (Zhou, 2011[51]).

A similar approach to foreign language learning is the Japanese method yakudoku. This dates back more than 1,000 years when Chinese was first studied in Japan as a written language (Hino, 1988[52]), relying on the translation of texts from the target language word by word into the students' first language. Yakudoku is still used to some extent in Japanese schools (Gorsuch, 2001[53]; Hosoki, 2011[54]; Kern, 2000[55]), but it is not the only teaching method historically used in Japan. For example, in the Edo period (1603 -1868), spoken Chinese was studied and taught by Tōtsūji (Chinese translators) and their descendants at the Nagasaki port for international trade between Japan and China and other countries of South Asia (Tiezheng, 2018[56]).

In the Western world, with the rise of vernacular languages (e.g. English, French, German, and Italian) and the diminishing importance of Latin as a lingua franca, the focus of Latin teaching also shifted from utilitarian to the analysis of grammar and rhetoric of classical Latin. "Modern languages" started to be taught in European schools in the 18th century (Richards and Rodgers, 2014[57]) using the "grammar translation" approach. This approach followed the same basic principles being used for Latin teaching (Richards and Rodgers, 2014[57]), focusing on grammar and reading, rather than speaking, and translating out of and into the foreign language, the latter of which was a novel feature (Howatt and Widdowson, 2004[58]). The drawback of this method is that while students developed a good understanding of grammar, they typically had more difficulties in using the languages for communicative purposes. Approaches focusing on grammar and translation continue to be used in some form in many parts of the world today, particularly when studying literature, grammar and vocabulary (Richards and Rodgers, 2014[57]).

19th and 20th century

The 19th and 20th century saw the rise of many approaches for teaching foreign languages. Often new approaches were developed outside formal schooling, for example in private language schools but also in girls' education, before being adopted more widely in education systems (McLelland, 2017[36]).

The "direct method" was based on the principles of first language acquisition, presenting the new language in contexts that explain its meaning, and teachers using only the target language (Richards and Rodgers, 2014[57]; Celce-Murcia, 2014[49]). The obvious drawback of this method is that teachers need to be fluent speakers of the target language, which can be an obstacle to implementing this approach in some contexts (Celce-Murcia, 2014[49]). The method became very popular in some parts of Europe and was also briefly experimented with in China, where it existed alongside other teaching approaches for a while (though it did not leave a lasting impact on language teaching (Hu, 2002[59])).

The "reading approach" was introduced in the United States. It focused on developing students' reading skills by introducing them to major literary works in the target language (Celce-Murcia, 2014[49]).

The "audio-lingual" and "oral-situational" approaches both emphasise oral skills. The "audio-lingual" approach developed out of the Army Specialised Training Program in World War II, which trained military personnel in the United States to attain conversational fluency in a range of languages quickly. The approach was founded on behaviourism and the assumption that language learning is based on the repetition of patterns until they become habits (Celce-Murcia, 2014[49]; Richards and Rodgers, 2014[57]). While relatively successful in a high-input setting, the method was found to be less successful when applied to typical classroom settings (Richards and Rodgers, 2014[57]).

In China the audio-lingual method co-existed with more grammar-focused approaches and, to a lesser extent, the direct method (Hu, 2002[59]). It took a slightly different form than in the United States, with more emphasis on the written form and the reading of classics (Yu, 2016[60]). The "oral-situational" approach, developed in Britain, organised language learning around situations (e.g. at the restaurant, at the pharmacy). All language material was presented in oral form before writing, and new grammatical and lexical items were introduced in the context of situations (Celce-Murcia, 2014[49]).

The second half of the 20th century saw a surge in the development of new teaching approaches, for example the "cognitive", "affective humanistic", "natural approach", "full body response", "silent way", or "suggestopedia" approaches (e.g. (Richards and Rodgers, 2014[57]; Celce-Murcia, 2014[49])).

Current approaches

Communicative language teaching (CLT) marked a paradigm shift in the 20th century (Richards and Rodgers, 2014[57]). This approach is based on an understanding of language as communication and learning is arranged around situations with the aim of developing communicative competence across the four skills of reading, writing, speaking and listening (see PISA 2025 Foreign Language Assessment Background Questionnaires Framework 2.8.1). Some authors have distinguished between strong and weak versions of the communicative approach, where the strong version develops language through communication whereas the weak version gives priority to providing the learner with opportunities to speak the target language (Howatt, 1984, p. 279[61]). In curricula and textbooks, the weak version usually predominates.

While CLT has evolved into quite diverse teaching practices, for example the Natural Approach, Content- or Task-Based Teaching (Richards and Rodgers, 2014[57]), the general principles of communicative language teaching continue to apply and are widely accepted today and incorporated in curricula all over the world as the predominant method in foreign language teaching (Butler and Iino, 2005[62]; Choi, 2007[63]; Wak Kam, 2002[64]; Hu, 2002[59]; Criado and Sanchez, 2009[65]; Richards and Rodgers, 2014[57]).

The definition of (foreign) language ability in terms of can-do statements and of communicative functions, as in the CEFR, reflects this paradigm shift and has contributed to the success of this approach.

Though CLT enjoys wide support in the research and teaching community as an approach to develop communication skills, there are conceptual, classroom-level and societal-institutional constraints that challenge its applicability and effectiveness in some contexts. Examples of these constraints are the lack of teacher training, the large size of classes, contradictions between the CLT approach and the expected teacher role, and lack of alignment with the assessment systems that are in place (Mutekwa, 2013[66]; Goto Butler, 2011[67]; Hasanova and Shadieva, 2008[68]; Gorsuch, 2001[53]; Littlewood, 2007[69]; Rao, 1996[70]).

Current developments

The research community is still actively debating what foreign language skills and knowledge learners need to acquire, and how they can do so. Examples of issues currently discussed amongst researchers and practitioners include, among others:

- the skills and knowledge needed to interact most effectively in foreign languages with new technologies (e.g. “digital translanguaging” (Schreiber, 2015[71]); “netspeak” (Tong, 2019[72]; Miola, 2013[73]))
- the trend towards more autonomous learning (Nunan and Richards, 2015[74])
- the importance of plurilingualism and mediation and related pedagogical approaches (Stathopoulou, 2015[75]; Heugh, 2018[76]; Jenkins, 2017[77]; Council of Europe, 2020[16])
- the reassessment of native-speaker norms, particularly for languages that function as a lingua franca (Council of Europe, 2018[78]; Seidlhofer, 2017[79]).

It remains to be seen how these developments will affect the teaching and learning of languages.

The need for data

Governments are keen to know if their efforts to promote foreign language teaching and learning are effective. They would like to compare their students’ competencies in foreign languages with their own policy goals, international standards and other countries’ performance. Policy makers want to know how they can improve their students’ performance and learn about the most effective policies, and teaching approaches and methods from the countries that are most successful in providing young people with the foreign language skills they will need in the future – and right now.

Educators and the public at large are also interested in international comparisons. This interest is reflected, for example, in the use of rankings based on results from non-representative free online tests (such as the EF English Proficiency Index (Education First, 2018[80])) and the students’ IELTS or TOEFL test results (see Ye (2016[81]) and Hongo (2013[82])). These indices can influence public opinion but, since they are not representative of the population of young people, can provide a biased picture of foreign language proficiency.

The strong policy interest in foreign language learning and teaching has already prompted a number of international studies, as described below (Box 1.2). However, as of this writing, there are no plans to conduct an international foreign language-skills assessment – except for PISA’s plan.

Box 1.2. International language assessments in the past

In the early 1970s, the International Association for the Evaluation of Educational Achievement (IEA) incorporated English and French foreign language assessments as part of its Six Subject Survey. Eight countries conducted the French survey: Chile, England, the Netherlands, New Zealand, Romania, Scotland, Sweden and the United States. Ten countries/economies conducted the English survey: Belgium (French Community), Chile, Finland, Germany (FRG), Hungary, Israel, Italy, the Netherlands, Sweden and Thailand. The surveys assessed 14-year-old students, and students in their final year of lower secondary school, in reading, speaking, listening and writing. The data were collected from 1970 to 1971 (Cumming, 1996[83]).

Two decades later, the IEA planned a three-phase project to assess foreign language skills. The assessment targeted 15-16 year-old students at the end of compulsory schooling, and students who were completing upper secondary school. The participating countries/economies included: Austria, Cyprus[1], the Czech Republic, Denmark, England, Finland, France, Hong Kong, Hungary, Iran, Israel, Italy, Latvia, the Netherlands, Norway, Philippines, Portugal, the Russian Federation, Slovenia, South Africa, Spain, Sweden, Switzerland, Thailand and the United States. However, due to a lack of funding, the IEA only carried out the first of the three planned phases. This phase focused on the national profiles of language education, describing curricula, policies, learning outside of school and teacher characteristics. In 1995, data were collected from 25 education systems on four commonly taught foreign languages: English, French, German and Spanish (IEA, 2019[84]; Cumming, 1996[83]).

Between 2010 and 2011, the European Commission implemented the first European Survey on Language Competences (European Commission, 2012[85]). Fourteen countries and economies participated: Belgium (the Flemish, French and German Communities), Bulgaria, Croatia, England, Estonia, France, Greece, Malta, the Netherlands, Poland, Portugal, Slovenia, Spain and Sweden. It assessed reading, listening and writing in English, French, German, Italian and Spanish. In every country, students were assessed in their first and second foreign languages. The first foreign language was English for most countries, except for England, and the Flemish and German Communities of Belgium, where the first foreign language was French. The participating students were either in their last year of lower secondary education (ISCED 2) or in their second year of upper secondary education (ISCED 3).

Measuring foreign language skills in PISA

As today's largest international assessment of student performance, PISA is well-placed to provide comparable data to answer questions about students' proficiency in a foreign language and on the background factors that are related to proficiency. The resulting analyses will answer questions about the school and national policies associated with foreign language teaching and learning. The wide range of data collected through the background questionnaires in PISA will further enrich the analyses of the language assessment, and provide input to direct political and economic efforts to improve teaching and learning.

As with the core PISA domains of reading, mathematics and science, the main focus of the foreign language assessment in PISA will be the proficiency of 15-year-old students in demonstrating and applying knowledge and skills. As in the other PISA domains, foreign language proficiency will be assessed using an instrument designed to provide data that are valid, reliable and interpretable.

A distinct advantage of PISA is the possibility it offers to compare foreign language proficiency with reading literacy in the students' language of instruction, which has been identified in the literature as an important factor associated with foreign language learning.

To guide the development of the cognitive instruments, the PISA Foreign Language Assessment developed a cognitive framework aligned with the Common European Framework of Reference for Languages (CEFR) and its new companion volume (Council of Europe, 2001[86]; 2020[16]), so that results can be interpreted according to the well-known language proficiency levels that are used in many countries around the world. The framework is also based on a socio-cognitive validation framework commonly used in foreign language assessments for test development and validation (Weir, 2005[87]; Shaw and Weir, 2007[88]; Khalifa and Weir, 2009[89]; Taylor, 2011[90]; Taylor and Geranpayeh, 2013[91]).

A framework for the background questionnaires that addresses factors related to governments and school systems, teachers' training and profile, teaching practices and students' background, was also developed.

The regular recurrence of PISA, and the benefits of working within an already established assessment, allow for an analysis of trends too. Thus, it will be possible to see how students' proficiency in foreign languages, school and teaching practices, foreign language curricula and contextual factors evolve over time.

Notes

1 The information in this document with reference to "Cyprus" relates to the southern part of the Island. There is no single authority representing both Turkish and Greek Cypriot people on the Island. Turkey recognises the Turkish Republic of Northern Cyprus (TRNC). Until a lasting and equitable solution is found within the context of the United Nations, Turkey shall preserve its position concerning the "Cyprus issue". The Republic of Cyprus is recognised by all members of the United Nations with the exception of Turkey. The information in this document relates to the area under the effective control of the Government of the Republic of Cyprus.

References

Ahuja, A., T. Chucherd and K. Pootrakool (2006), *Human Capital Policy: Building a Competitive Workforce for 21 st Century Thailand*, Bank of Thailand, Bangkok, https://www.bot.or.th/Thai/MonetaryPolicy/ArticleAndResearch/SymposiumDocument/paper5_2549.pdf. [21]

Ammon, U. (2015), *Die Stellung der Deutschen Sprache in der Welt. The Position of the German Language in the World*, De Gruyter, Berlin. [44]

Araújo, L. et al. (2015), *Languages and Employability*, Publications Office of the European Union, Luxembourg, http://dx.doi.org/10.2788/860807. [17]

Bada, F. (2018), *Languages Most Commonly Used On The Web*, WorldAtlas, https://www.worldatlas.com/articles/languages-most-commonly-used-by-the-web.html (accessed on 17 September 2020). [43]

Bak, T. et al. (2016), "Novelty, challenge, and practice: the impact of intensive language learning on attentional functions", *PLoS ONE*, Vol. 11/4, http://dx.doi.org/e0153485. [33]

Bamford, K. and D. Mizokawa (1991), "Additive-Bilingual (Immersion) Education: Cognitive and Language Development", *Language Learning*, Vol. 41/3, pp. 413-429, http://dx.doi.org/10.1111/j.1467-1770.1991.tb00612.x. [31]

Bokor, M. (2018), "English Dominance on the Internet", *The TESOL Encyclopedia of English Language Teaching*, pp. 1-6, http://dx.doi.org/10.1002/9781118784235.eelt0281. [41]

Butler, Y. and M. Iino (2005), "Current Japanese Reforms In English Language Education: The 2003 "Action Plan"", *Language Policy*, Vol. 4/1, pp. 25-45. [62]

Canadian Heritage (2016), *Economic Advantages of Bilingualism Literature Review*, Canadian Heritage, https://www.caslt.org/files/learn-languages/pch-bilingualism-lit-review-final-en.pdf. [18]

Celce-Murcia, M. (2014), "An Overview of Language Teaching Methods and Approaches", in Celce-Murcia, M., D. Brinton and M. Snow (eds.), *Teaching English as a Second or Foreign Language*, Heinle & Heinle, Boston, MA. [49]

Choi, Y. (2007), "The history and policy of English language education in Korea", in Choi, Y. and B. Spolsky (eds.), *English education in Asia: History and policies*. [63]

Colombian National Ministry for Education (2006), "Formar en Lenguas Extranjeras: Inglés!el Reto!", *Guías*, No. 22, Colombian National Ministry for Education, Bogotá, https://www.mineducacion.gov.co/1621/articles-115174_archivo_pdf.pdf. [40]

Council of Europe (2020), *Common European Framework of Reference for Languages : learning, teaching, assessment. Companion volume,*, Council of Europe Publishing, Strasbourg, http://www.coe.int/lang-cefr. [16]

Council of Europe (2018), *Common European Framework of Reference for Languages: Learning, Teaching, Assessment. Companion Volume with new descriptors*, Language Policy Programme Education Policy Division Education Department, http://www.coe.int/lang-cefr. [78]

Council of Europe (2001), *Common European Framework of Reference for Languages. Learning, Teaching, Assessment*, Cambridge University Press, Cambridge, http://www.coe.int/lang-CEFR. [86]

Council of the European Union (2002), *Barcelona European Council - Presidency conclusions*, Council of the European Union, Barcelona, http://ec.europa.eu/invest-in-research/pdf/download_en/barcelona_european_council.pdf (accessed on 22 February 2019). [47]

Criado, R. and A. Sanchez (2009), "English language teaching in Spain: Do textbooks comply with the official methodological regulations? A sample analysis", *International Journal of English Studies*, Vol. 9/1, pp. 1-28. [65]

Crystal, D. (1987), *The Cambridge Encyclopedia of Language*, Cambridge University Press, Cambridge. [39]

Cumming, A. (1996), "IEA's Studies of Language Education: Their Scope and Contributions", *Assessment in Education*, Vol. 3, pp. 179-192, https://doi.org/10.1080/0969594960030205. [83]

Curtain, H. and C. Dahlberg (2004), *Languages and children--making the match : new languages for young learners*, Allyn & Bacon. [11]

Della Chiesa, B., J. Scott and C. Hinton (2012), *Languages in a GlobalWorld: Learning for Better Cultural Understanding*, OECD Publishing, https://www.oecd-ilibrary.org/docserver/9789264123557-en.pdf?expires=1559551110&id=id&accname=ocid84004878&checksum=77AEA46C12225DA61CDF25B5B7389A43. [14]

Della Chiesa, B., J. Scott and C. Hinton (eds.) (2012), *Preface: Language Learning and Culture in a Time of Globalisation*, Languages in a Global World: Learning for Better Cultural Understanding, OECD Publishing, Paris, https://dx.doi.org/10.1787/9789264123557-en. [10]

Education First (2018), *EF English Proficiency Index*, EF, https://www.ef.com/__/~/media/centralefcom/epi/downloads/full-reports/v8/ef-epi-2018-english.pdf. [80]

Euromonitor International (2012), *The Benefits of the English Language for Individuals and Societies: Quantitative Indicators from Algeria, Egypt, Iraq, Jordan, Lebanon, Morocco, Tunisia and Yemen*, British Council. [25]

European Commission (2012), *First European Survey on Language Competences: Technical Report*, European Commission, Bruxelles. [85]

European Commission (2012), *Language Competences for Employability, Mobility and Growth*, European Commission, Strasbourg. [19]

European Commission (2010), *Lingua Franca: Chimera or Reality?*, Publications Office of the European Union. [45]

Fidrmuc, J. and J. Fidrmuc (2016), "Foreign languages and trade: evidence from a natural experiment", *Empirical Economics*, Vol. 50/1, pp. 31-49, http://dx.doi.org/10.1007/s00181-015-0999-7. [27]

Flemish Ministry for Education and Training (2000), *Derde graad - ASO - Moderne vreemde talen Frans-Engels*, https://www.kwalificatiesencurriculum.be/derde-graad-aso-moderne-vreemde-talen-frans-engels (accessed on 29 July 2019). [5]

Foreman-Peck, J. and Y. Wang (2014), *The Costs to the UK of Language Deficiencies as a Barrier to UK Engagement in Exporting: A Report to UK Trade & Investment*, Cardiff Business School, https://assets.publishing.service.gov.uk/government/uploads/system/uploads/attachment_data/file/309899/Costs_to_UK_of_language_deficiencies_as_barrier_to_UK_engagement_in_exporting.pdf (accessed on 4 October 2018). [28]

Garrouste, C. (2008), "Language Skills and Economic Returns", *Policy Futures in Education*, Vol. 6/2, pp. 187-202, http://dx.doi.org/10.2304/pfie.2008.6.2.187. [22]

Gorsuch, G. (2001), "Japanese EFL Teachers' Perceptions of Communicative, Audiolingual and Yakudoku Activities: The Plan Versus the Reality", *Educational Policy Archives*, Vol. 9/10. [53]

Goto Butler, Y. (2011), "The Implementation of Communicative and Task-Based Language Teaching in the Asia-Pacific Region", *Annual Review of Applied Linguistics*, Vol. 31, pp. 36-57. [67]

Grin, F. (1999), *Compétences et récompenses. La valeur des langues en Suisse*, Éditions universitaires, Fribourg. [23]

Gudykunst, W. (2003), *Cross-cultural and intercultural communication*, Sage Publications, https://uk.sagepub.com/en-gb/eur/cross-cultural-and-intercultural-communication/book226099 (accessed on 4 October 2018). [12]

Hasanova, D. and T. Shadieva (2008), "Implementing communicative language teaching in Uzbekistan", *TESOL Quaterly*, Vol. 42, http://dx.doi.org/10.1002/j.1545-7249.2008.tb00215.x. [68]

Heugh, K. (2018), "Conclusion: Multilingualism, Diversity and Equitable Learning", in Van Avermaet, P. et al. (eds.), *The multilingual edge of education*, Palgrave Macmillan, London, http://dx.doi.org/10.1057/978-1-137-54856-6. [76]

Hino, N. (1988), "Yakudoku: Japan's dominant tradition in foreign language learning", *JALT*, Vol. 10/1&2, pp. 45-53. [52]

Hongo, J. (2013), "Abe wants TOEFL to be key exam", *The Japan Times*, https://www.japantimes.co.jp/news/2013/03/25/national/abe-wants-toefl-to-be-key-exam/#.Xz-8wMgzbIU. [82]

Hosoki, Y. (2011), "English Language Education in Japan : Transitions and Challenges", *Memoirs of Kyushu International University*, Vol. 6, pp. 1-2. [54]

Howatt, A. (1984), *A History of English Language Teaching*, Oxford University Press, Oxford. [61]

Howatt, A. and H. Widdowson (2004), *A History of English Language Teaching*, Oxford University Press, Oxford. [58]

Hu, G. (2002), "English language teaching in the people's republic of China", in Silver, R., G. Hu and M. Iino (eds.), *English language education in China, Japan, and Singapore*, National Institute of Education, Nanyang Technological University. [59]

IEA (2019), *Other IEA Studies*, https://www.iea.nl/studies/iea/other#spy-para-160 (accessed on 30 July 2019). [84]

Isphording, I. (2015), "Language and Labor Market Success", in *International Encyclopedia of the Social & Behavioral Sciences*, Elsevier, http://dx.doi.org/10.1016/b978-0-08-097086-8.94033-9. [20]

Jenkins, J. (2017), "Not English but English-within-Multilingualism", in Coffey, S. and U. Wingate (eds.), *New Directions for Research in Foreign Language Education*, Routledge, New York, http://dx.doi.org/10.4324/9781315561561-5. [77]

Kern, R. (2000), *Literacy and Language Teaching*, Oxford University Press, Oxford. [55]

Khalifa, H. and C. Weir (2009), *Examining Reading: Research and Practice in Assessing Second Language Learning*, UCLES/Cambridge University Press, Cambridge. [89]

Kim, M. et al. (2014), "Lingua Mercatoria: Language and Foreign Direct Investment", *International Studies Quarterly*, Vol. 59/2, pp. 330-343, http://dx.doi.org/10.1111/isqu.12158. [29]

Littlewood, W. (2007), "Communicative and taskbased language teaching in East Asian classrooms", *Language Teaching*, Vol. 40, pp. 243-249, http://dx.doi.org/doi:10.1017/S0261444807004363. [69]

Marian, V. and A. Shook (2012), "The Cognitive Benefits of Being Bilingual", *Cerebrum: the Dana Forum on Brain Science*, Vol. 13, https://www.ncbi.nlm.nih.gov/pmc/articles/PMC3583091/ (accessed on 5 October 2018). [13]

McLelland, N. (2017), *Teaching and Learning Foreign Languages*, Routledge, Milton Park, Abingdon, Oxon; New York, NY : Routledge, [2017], http://dx.doi.org/10.4324/9781315624853. [36]

Melitz, J. (2008), "Language and foreign trade", *European Economic Review*, Vol. 52/4, pp. 667-699, http://dx.doi.org/10.1016/J.EUROECOREV.2007.05.002. [26]

MEXT (2017), *Shougakkou gakushuu shidou youryou kaisetsu: Gaikokugo Katsudou Gaikokugo Hen [Primary education curriculum guideline: Foreign languages]*, Japanese Ministry of Education, Culture, Sports, Science and Technology (MEXT), http://www.mext.go.jp/component/a_menu/education/micro_detail/__icsFiles/afieldfile/2019/03/18/1387017_011.pdf. [7]

Ministry of Education - New Zealand (2007), *The New Zealand Curriculum / Kia ora 2007 - NZ Curriculum Online*, http://nzcurriculum.tki.org.nz/The-New-Zealand-Curriculum#collapsible7 (accessed on 5 October 2018). [8]

Miniwatts Marketing Group (2019), *Internet world users by language: Top 10 languages*, https://www.internetworldstats.com/stats7.htm (accessed on 17 September 2020). [42]

Miola, E. (ed.) (2013), *Languages Go Web: Standard and Non-standard Languages on the Internet*, Edizioni dall'Orso, Alessandria. [73]

Murphy, V. et al. (2015), "The influence of learning a second language in primary school on developing first language literacy skills", *Applied Psycholinguistics*, Vol. 36, pp. 1133-1153. [34]

Musumeci, D. (2009), "A History of Language Teaching", in Long, D. and C. Doughty (eds.), *The Handbook of Language Teaching*, Blackwell Publishing, Chichester. [50]

Mutekwa, A. (2013), "The challenges of using the Communicative Approach (CA) in the teaching of English as a Second language (ESL) in Zimbabwe: Implications for ESL teacher education", *Africa Education Review*, Vol. 10/3, pp. 539-553, http://dx.doi.org/10.1080/18146627.2013.853547. [66]

Nunan, D. and J. Richards (eds.) (2015), *Language Learning Beyond the Classroom*, Routledge, London, http://dx.doi.org/prova123. [74]

OECD (2018), *Education at a Glance 2018: OECD Indicators*, OECD Publishing, Paris, https://dx.doi.org/10.1787/eag-2018-en. [1]

OECD (2018), *International Migration Outlook 2018*, OECD Publishing, Paris, https://dx.doi.org/10.1787/migr_outlook-2018-en. [2]

Porto, M., S. Houghton and M. Byram (2018), "Intercultural Citizenship in the (Foreign) Language Classroom", *Language Teaching Research*, Vol. 22/5, pp. 484-498, http://dx.doi.org/10.1177/1362168817718580. [15]

Rao, Z. (1996), "Reconciling Communicative Approaches to the Teaching of English with Traditional Chinese Methods", *Research in the Teaching of English*, Vol. 30/4, pp. 458-471. [70]

República de Costa Rica (2016), *Programas de Estudio de Ingles. Tercer Ciclo de la Educación General Básica y Educación Diversificada*, Ministerio de Educación Pública, San José. [6]

Richards, J. and T. Rodgers (2014), *Approaches and Methods in Language Teaching*, Cambridge University Press, Cambridge. [57]

Rothman, J., J. Cabrelli and K. De Bot (2013), "Third Language Acquisition", in Herschensohn, J. and M. Young-Scholten (eds.), *The Cambridge Handbook of Second Language Acquisition*, Cambridge University Press, Cambridge. [35]

Saiz, A. and E. Zoido (2002), "The returns to speaking a second language"*, 16*, No. 2, Federal Reserve Bank of Philadelphia, Philadelphia, https://www.phil.frb.org/-/media/research-and-data/publications/working-papers/2002/wp02-16.pdf. [24]

Schreiber, B. (2015), ""I am what I am": Multilingual Identity and Digital Translanguaging", *Language Learning & Technology*, Vol. 19/3, pp. 69-87, http://dx.doi.org/10125/44434. [71]

Seidlhofer, B. (2017), "English as a Lingua Franca: Why Is It So Controversial?", *JACET International Convention Selected Papers*, Vol. 5, pp. 2-24. [79]

Seidlhofer, B. (2011), *Understanding English as a Lingua Franca*, Oxford University Press, Oxford. [37]

Shaw, S. and C. Weir (2007), *Examining Writing: Research and Practice in Assessing Second Language Writing*, UCLES/Cambridge University Press, Cambridge. [88]

Stathopoulou, M. (2015), *Cross-Language Mediation in Foreign Language Teaching and Testing*, Multilingual Matters. [75]

Taylor, L. (2011), *Examining Speaking: Research and Practice in Assessing Second Language Speaking*, UCLES/Cambridge University Press, Cambridge. [90]

Taylor, L. and A. Geranpayeh (2013), *Examining Listening: Research and Practice in Assessing Second Language Listening*, UCLES/Cambridge University Press, Cambridge. [91]

Tenzer, H., S. Terjesen and A. Harzing (2017), "Language in International Business: A Review and Agenda for Future Research", *Management International Review*, Vol. 57/6, pp. 815-854, http://dx.doi.org/10.1007/s11575-017-0319-x. [30]

Tiezheng, Y. (2018), "Chinese Language Educators in Meiji-Era Japan: Miyajima Daihachi and Zhang Tingyan", in McLelland, N. and R. Smith (eds.), *The History of Language Learning and Teaching - Vol. III Across Cultures*, Legenda, Oxford. [56]

Tinsley, T. and K. Board (2017), *Language for the Future: The languages the United Kingdom needs to become a truly global nation*, British Council, https://wales.britishcouncil.org/sites/default/files/languages_for_the_future_report_1.pdf. [46]

Tong, L. (2019), *An Analysis on the Forms and Characteristics of English Netspeak*, Atlantis Press, http://dx.doi.org/10.2991/icssed-19.2019.89. [72]

UK Department for Education (2014), *The National Curriculum in England*, UK Government, London, http://www.gov.uk/dfe/nationalcurriculum. [9]

UNESCO (1953), *The Use of Vernacular Languages in Education*, https://unesdoc.unesco.org/ark:/48223/pf0000002897. [38]

UNWTO (2019), *World Tourism Barometer Vol. 17*, United Nations World Tourism Organization, Madrid, http://marketintelligence.unwto.org/content/unwto-world-tourism-barometer. [3]

Wak Kam, H. (2002), "English Language Teaching in East Asia Today: An Overview", *Asia Pacific Journal of Education*, Vol. 22/2, pp. 1-22, http://dx.doi.org/10.1080/0218879020220203. [64]

Weir, C. (2005), *Language Testing and Validation: An Evidence- Based Approach*, Palgrave McMillan, Hampshire. [87]

Widdowson, H. (2004), "A perspective on recent trends", in Howatt, A. and H. Widdowson (eds.), *A History of English Language Teaching*, Oxford University Press, Oxford. [48]

Woll, B. and L. Wei (2019), *Cognitive Benefits of Language Learning: Broadening our perspectives. Final Report to the British Academy*, British Academy, London, https://www.thebritishacademy.ac.uk/sites/default/files/Cognitive-Benefits-Language-Learning-Final-Report.pdf. [32]

World Bank (2019), *Trade (% of GDP)*, https://data.worldbank.org/indicator/NE.TRD.GNFS.ZS (accessed on 29 July 2019). [4]

Ye, J. (2016), "Study ranks Hong Kong third in Asia for English, trailing Malaysia and Philippines | South China Morning Post", *South China Morning Post*, https://www.scmp.com/news/hong-kong/education-community/article/2019271/study-ranks-hong-kong-no-3-asia-english-trailing (accessed on 12 October 2018). [81]

Yu, X. (2016), "Domestication of Audiolingualism in ELT in China: From the Perspective of Cultural Accommodation.", *Studies in Literature and Language*, Vol. 13/3, pp. 22-27. [60]

Zhou, W. (2011), "L'enseignement des langues étrangères en Chine ancienne", *Synergies Chine*, Vol. 6, pp. 189-194. [51]

2 Defining foreign language proficiency

This chapter provides the definition of foreign language that will be used in the PISA Foreign Language Assessment. This definition will be used to guide instrument development and to define the target population. The chapter also defines what it means to be proficient in a foreign language in the context of this assessment. It then introduces two existing frameworks that the assessment will build on: the Common European Framework of Reference for Languages (CEFR) and the socio-cognitive model of language use based on Weir's 2005 model of test development and validation and further developed by other authors for the different foreign language skills.

Foreign language

There is no universal agreement on the definition of a foreign language. This section provides a brief overview of different definitions of foreign language, and presents the definition that the Programme for International Student Assessment (PISA) will use for its Foreign Language Assessment (FLA) component.

A first distinction to consider is between a first language and a foreign or second language. A first language (L1), also sometimes referred to as native language or mother tongue, is generally considered to be the first language a person acquires during early childhood and before any other languages learned (Harley, 2014[1]). Some children grow up hearing more than one language from birth, which is generally referred to as "simultaneous bilingualism" (Harley, 2014[1]) or "bilingual first language acquisition" (De Houwer, 2009[2]).

Typically, the boundary for first language acquisition is considered to be the age of three years, i.e. after the grammar of the first language has been mostly established (Lakshmanan, 2009[3]). Any languages learned afterwards would be considered second or foreign languages. However, some argue that the boundaries between first and second language acquisition are not as clear cut (Foster-Cohen, 2001[4]). The role of age in language acquisition is widely researched and strongly debated ((Birdsong, 2006[5]; Singleton and Ryan, 2004[6]; Muñoz and Singleton, 2011[7]).

A second distinction to consider is between a "foreign" and a "second" language, which forms part of an ongoing debate in the field of applied linguistics. Some linguistic theories suggest that the main difference between the two lies in the setting in which language learning takes place. From this perspective, foreign language refers to a language that generally has no direct link with the learner's immediate environment (Yule, 2010[8]; Punchihetti, 2013[9]). A second language, in contrast, is a language that the learner can easily encounter in daily life. Thus, learners of a second language have more opportunities to interact with the target language in their immediate environment than foreign language learners (Yule, 2010[8]; Moeller and Catalano, 2015[10]). However, in the present globalised era, the distinction between foreign and second language is not always clear. For example, situations in which students interact with individuals who do not speak the students' native language are becoming more frequent, and mass media provide increasing opportunities for students to practice their language skills (Ipek, 2009[11]).

Some argue that this distinction can also serve to distinguish content and language integrated learning (CLIL) from immersion programmes, where CLIL programmes tend to teach a language that is foreign to students in the sense that they do not encounter it in everyday life, while the language used in immersion programmes is present in students' environment (Lasagabaster and Sierra, 2010[12]; Dalton-Puffer, Nikula and Smit, 2010[13]). However, this distinction is not always clear cut and is sometimes debated (Eurydice, 2006[14]; Cenoz, Genesee and Gorter, 2014[15]; Nikula, Dalton-Puffer and Llinares, 2013[16]).

Some linguistic disciplines do not distinguish between a foreign and a second language (Littlewood, 1984[17]; Ellis, 2015[18]; Lightbown and Spada, 2013[19]) and use the term "second language" to refer to all languages learned after the first one (Gass and Selinker, 2008[20]; Long, 1993[21]; Mehisto, Frigols and Marsh, 2008[22]; Yule, 2010[8]; Ellis, 2015[18]).

Another distinction concerns language learning and language acquisition, where language learning refers to a more conscious process that takes place in an institutional and instructional setting, while language acquisition refers to a gradual development of ability that emerges naturally by using the language in communicative situations (Crystal, 2008[23]; Yule, 2010[8]). However, the relevance of this distinction is sometimes debated, based on the argument that language learning and acquisition are processes that can take place in both types of settings (Ellis, 2015[18]). Some suggest the term "language development" as a way to better reflect that linguistic skills do not necessarily develop linearly but can grow and decline with time (De Bot and Larsen-Freeman, 2011[24]).

Whether ancient languages should be considered in the same way as modern foreign languages is also debated, and school curricula often make this distinction, following the argument that they pursue different goals. The main goal for modern language learning is to develop communicative competence, i.e. the ability to interact with other speakers of the language in appropriate ways in social, educational or professional contexts. In contrast, ancient language teaching traditionally focuses on linguistic analysis, translation, composition and reading as a means to foster understanding of the ancient language itself as well as modern languages and ancient cultures (Carpenter, 2000[25]; Carlon, 2013[26]). These differing goals result in distinct assessment criteria (e.g. International Baccalaureate [IB], American Council on the Teaching of Foreign Languages [ACTFL]) and different prevalent teaching methods. However, historically, modern language teaching methods were derived from ancient language teaching and some of these methods continue to be influential today (e.g. grammar-translation methods). There are also recent movements that seek to apply second language acquisition research to Latin and Greek teaching (Carlon, 2013[26]; Koutropoulos, 2011[27]; Dugdale, 2011[28]; Patrick, 2019[29]).

Some organisations use broader terms that aim to encompass more nuances of language learning, showing a development towards a more integrated approach to language learning across the curriculum that includes foreign/second languages, the language/s of schooling and students' home languages. For example, the European Centre for Modern Languages of the Council of Europe uses the term "language learning/education" to go beyond foreign language teaching to incorporate all aspects of schooling including the majority school language, home languages and the languages needed to learn subjects (ECML, n.d.[30]).

Similarly, the Common European Framework of Reference for Languages (CEFR), a policy instrument developed by the Council of Europe, refers to "foreign language/second language", and its updated descriptors focus on the learning of modern foreign languages, but also reach beyond the area of modern language learning to include elements of language education across the curriculum (Council of Europe, 2020[31]). The functional component of the CEFR's description of language use applies in principle to all languages, including the user/learner's first language.

For assessment and research purposes, an operational definition is required. For example, for its system-level questionnaires, the Eurydice network, established by the European Commission to provide information on education systems and policies, uses an education-related definition and defines foreign language as a language that is foreign, modern, second (or third) in the curriculum laid down by the central (or top-level) education authorities (European Commission/EACEA/Eurydice, 2017, p. 142[32]). As this definition is unrelated to the political status of a language, certain languages regarded as regional or minority languages may be included in the curriculum of some countries as foreign languages. Classical languages may be included or not, depending on whether they are included under the umbrella term of "foreign languages" or seen as separate from "modern languages" in a given education system.

The PISA Foreign Language Assessment will use a definition of foreign language that is both operational and education-related; it will be used to define the target population, clarifying which countries this assessment is appropriate for, and which students are eligible to take the assessment; and it will guide instrument development. This assessment would be appropriate to assess any modern language formally taught in school settings other than the main language of instruction. The assessed languages will most likely not be official languages of the country, but in some cases, they could be, particularly in multilingual countries. While the assessed language can be one of the languages of instruction, it should not be the main language of instruction (which is normally used, in the PISA context, for the assessment of reading). However, it would be relevant to make an exception in schools where, due to a system-level policy to improve students' language skills, the main language of instruction is not the language spoken by students at home and in their social environment. For instrument development, foreign languages will be understood as all modern languages that are formally taught in school settings, other than the main language of schooling, which is defined as the language of the PISA reading test.

Foreign language proficiency

The PISA Foreign Language Assessment aims to make the most of extensive international experience in assessing foreign languages and build on prior foreign language studies and pre-existing foreign language competence frameworks. It builds on two frameworks in particular.

The first framework is the CEFR descriptive scheme (Council of Europe, 2001[33]; Council of Europe, 2020[31]), an internationally recognised framework for describing foreign language proficiency that offers a broad, comprehensive overview of foreign language performance at different levels of ability. The scheme lists and describes the general competences, language competences, activities and strategies involved in successful communicative foreign language use from a beginner's level up to a level of highest attainment in language proficiency, and identifies six key proficiency levels (labelled A1, A2, B1, B2, C1 and C2), which can be expanded to Pre-A1 level, and have sublevels, e.g. A1.1, A2.1. The six CEFR language proficiency levels are widely recognised and used in many foreign language settings and educational systems around the world. Its 2001 publication, for instance, has been translated into over 40 languages (Council of Europe, 2020[31]), and the development of the revised and supplementary CEFR descriptors in the 2018 edition was piloted in educational institutes from 26 countries geographically outside of Europe (out of a total of 58 total participating countries in the pilot).

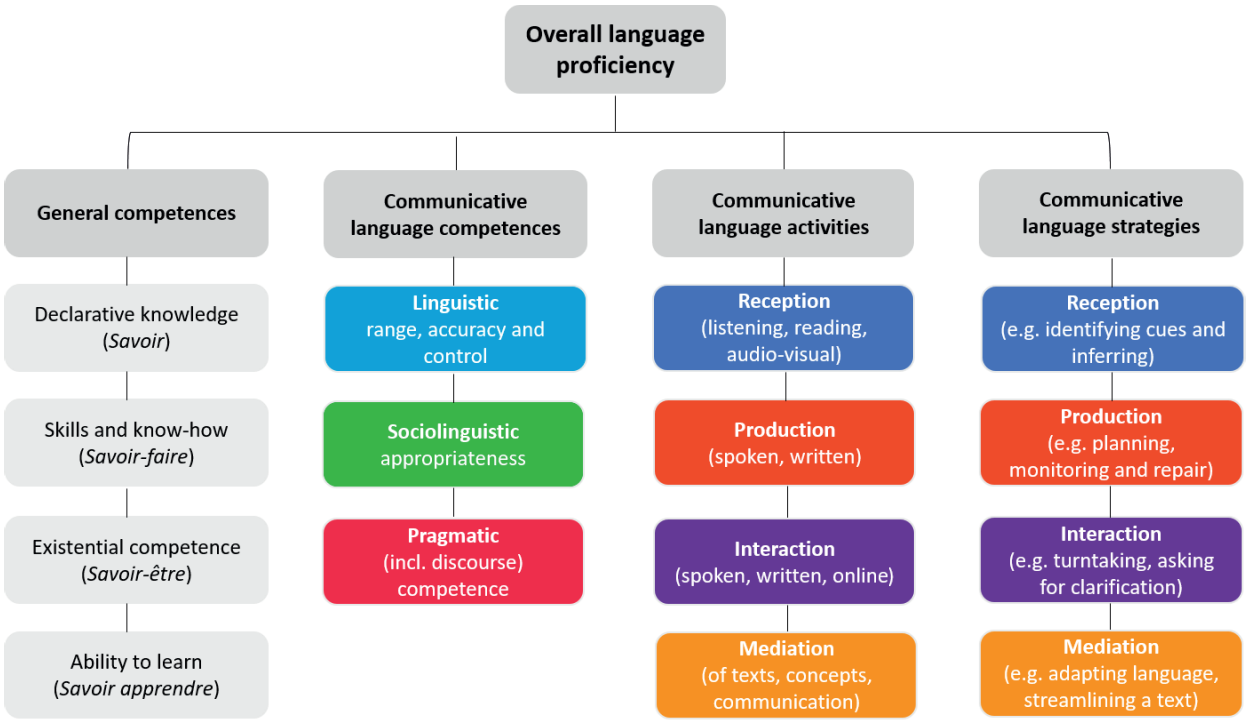
The educational value of the CEFR is not confined to its well-recognised proficiency benchmark levels, however. One of its prime aims is to promote quality in foreign language learning, teaching and assessment:

> *"… the CEFR broadens the perspective of language education in a number of ways, not least by its vision of the user/learner as a social agent, co-constructing meaning in interaction, and by the notions of mediation and plurilingual/pluricultural competences. The CEFR has proved successful precisely because it encompasses educational values, a clear model of language-related competences and language use, and practical tools, in the form of illustrative descriptors, to facilitate the development of curricula and orientation of teaching and learning." (Council of Europe, 2020, p. 21[31])*

The second framework is a socio-cognitive model of language use based on Weir's (2005[34]) model of test development and validation, and its extension and implementation in the context of the four language skills by Shaw and Weir (2007[35]) for writing; by Khalifa and Weir (2009[36]) for reading; by Taylor (2011[37]) for speaking; and by Taylor and Geranpayeh (2013[38]) for listening. It is a well-established framework and has been applied to a wide variety of language assessment contexts around the world. Selected examples of the framework are applied in different contexts, including in Asia, North Africa, Europe and Latin America (Ahmad and Abidin, 2020[39]; Bannur, Abidin and Jamil, 2015[40]; Boyd and Taylor, 2016[41]; Dunlea, 2015[42]; Florescano et al., 2011[43]; Liu and Pan, 2019[44]; Oliveri, 2019[45]; Selvaruby, O'Sullivan and Watts, 2008[46]; Su, Weir and Wu, 2019[47]; Tozlu and Ünaldı, 2018[48]). It has also been cited in validation documents for international language tests (e.g. Cambridge English Qualifications, Aptis).

Both frameworks espouse a socio-cognitive approach to language description and use. The CEFR's description of foreign language proficiency is often described as "socio-cognitive" in the sense that it describes communicative language activities ("socio") and communicative language competences ("cognitive") (see Figure 2.1). Likewise, the socio-cognitive model views language use as an interplay of social factors (the linguistic and task[1] demands implicated in successful language use) and cognitive factors (the cognitive processes employed for successful language use). While the CEFR is primarily a foreign language descriptive scheme, the socio-cognitive model is a language test validation model used here to complement the CEFR's description by analysing the cognitive processes that support the distinct language skills (reading, listening, speaking, writing), highlighting the role of meta-cognitive regulation strategies, and describing how context and task features, including linguistic ones, activate (or inhibit) particular cognitive processes.

Figure 2.1. The CEFR descriptive scheme for language proficiency



Note: Additional information has been included on most boxes to provide further clarification on the categories. For further information on what has been added to the original figure, see Annex 2.A.1.
Source: Adapted from Council of Europe (2020[31]), *Common European Framework of Reference for Languages: Learning, teaching, assessment - Companion Volume*, © Council of Europe, Strasbourg, p. 32.

The PISA definition of foreign language proficiency (Box 2.1) draws heavily on the CEFR description (Council of Europe, 2001, pp. 9-10[33]; Council of Europe, 2020, pp. 31-33[31]), which is schematically depicted in Figure 2.1. Each part of the definition is discussed in the remainder of this section.

Box 2.1. The PISA definition of foreign language proficiency

Foreign language proficiency is the ability to use a foreign language to communicate effectively. It requires a combination of communicative language competences and general competences that allow the foreign language learner to perform communicative language activities (reception, production, interaction and mediation), which involve one or a combination of the following skills: reading, listening, speaking or writing. It also requires the activation of appropriate language strategies.

In alignment with the communicative approach and the CEFR, the PISA definition of foreign language proficiency encompasses all four modes of communication (reception, production, interaction, mediation) and fully acknowledges that they are inseparable elements of communicative foreign language use. However, due to practical and logistical constraints, the PISA 2025 Foreign Language Assessment will be organised by skill[2] (reading, listening, speaking and writing), and the 2025 cycle will focus on reception (reading and listening comprehension) and production (speaking) only. The additional demands of testing and marking spoken interaction, writing or mediation, combined with the high numbers of PISA test takers, currently render the assessment of these modes of communication unfeasible. The PISA Programme

intends to explore the assessment of the remaining modes of communication in future iterations of the FLA.

Effective foreign language use requires communicative language competences …

Communicative language competences are subdivided into linguistic, sociolinguistic and pragmatic competences.[3] The CEFR offers illustrative descriptors that outline how these competences evolve from beginner (pre-A1, A1) to mastery level (C2) (Council of Europe, 2020, pp. 130-142[31]; Council of Europe, 2018, pp. 131-144[49]). For example:

- **Linguistic competence** captures lexical, syntactic, phonological, phonetic and orthographic dimensions of knowledge and skills of language as a system. A distinction is drawn between *range* of linguistic competence, with scales covering "general linguistic range" and "vocabulary range", and *control*, with scales describing "grammatical accuracy", "vocabulary control", "phonological control" and "orthographic control". This distinction is intended to draw attention not only to accuracy (control) but also to the complexity and breadth (range) of the linguistic means used by foreign language learners.
- **Sociolinguistic competence** refers to the knowledge and skills necessary for successful communication within a social context (e.g. awareness of politeness conventions, norms of socialising, sociocultural cues, use of appropriate register, etc., as per the "sociolinguistic appropriateness" scale).
- **Pragmatic competence** is defined as "knowledge of the principles of language use according to which messages are: 1) organised, structured and arranged ("discourse competence"); 2) used to perform communicative functions ("functional competence"); and 3) sequenced according to interactional and transactional schemata ("design competence")" (Council of Europe, 2020, p. 137[31]).

… and general competences …

General competences are more closely associated with factors within individuals in terms of their knowledge of the world and their ability and/or willingness to learn. As such, they are non-linguistic factors, and will not, therefore, be assessed in the PISA FLA test. However, they are briefly considered here as they can influence (facilitate or hinder) understanding and communication in a foreign language because they typically relate to how individuals integrate new knowledge into their existing knowledge base and make sense of similarities or differences between the foreign language view of the world and their own. The PISA FLA will control for such influences (see Chapter 4).

The **general competences** comprise four components:

- **Declarative knowledge (*savoir*)** is "knowledge of the world", resulting from experience (empirical knowledge) and more formal learning (academic knowledge).
- **Skills and know-how (*savoir-faire*)** are about the ability to carry out procedures, e.g. knowing how to drive a car.
- **Existential competence (*savoir-être*)** is "the sum of the individual characteristics, personality traits and attitudes which concern, for example, self-image and one's view of others and willingness to engage with other people in social interaction" (Council of Europe, 2001, p. 12[33]). Existential competence can be modified and is considered an aspect of an individual's general abilities.
- The **ability to learn (*savoir apprendre*)** is "knowing how, or being disposed, to discover 'otherness' – whether the other is another language, another culture, other people or new areas of knowledge" (Council of Europe, 2001, p. 13[33]). The ability to learn is particularly relevant to language learning.

… which allow the foreign language user/learner to perform communicative language activities … they also require the activation of appropriate language strategies

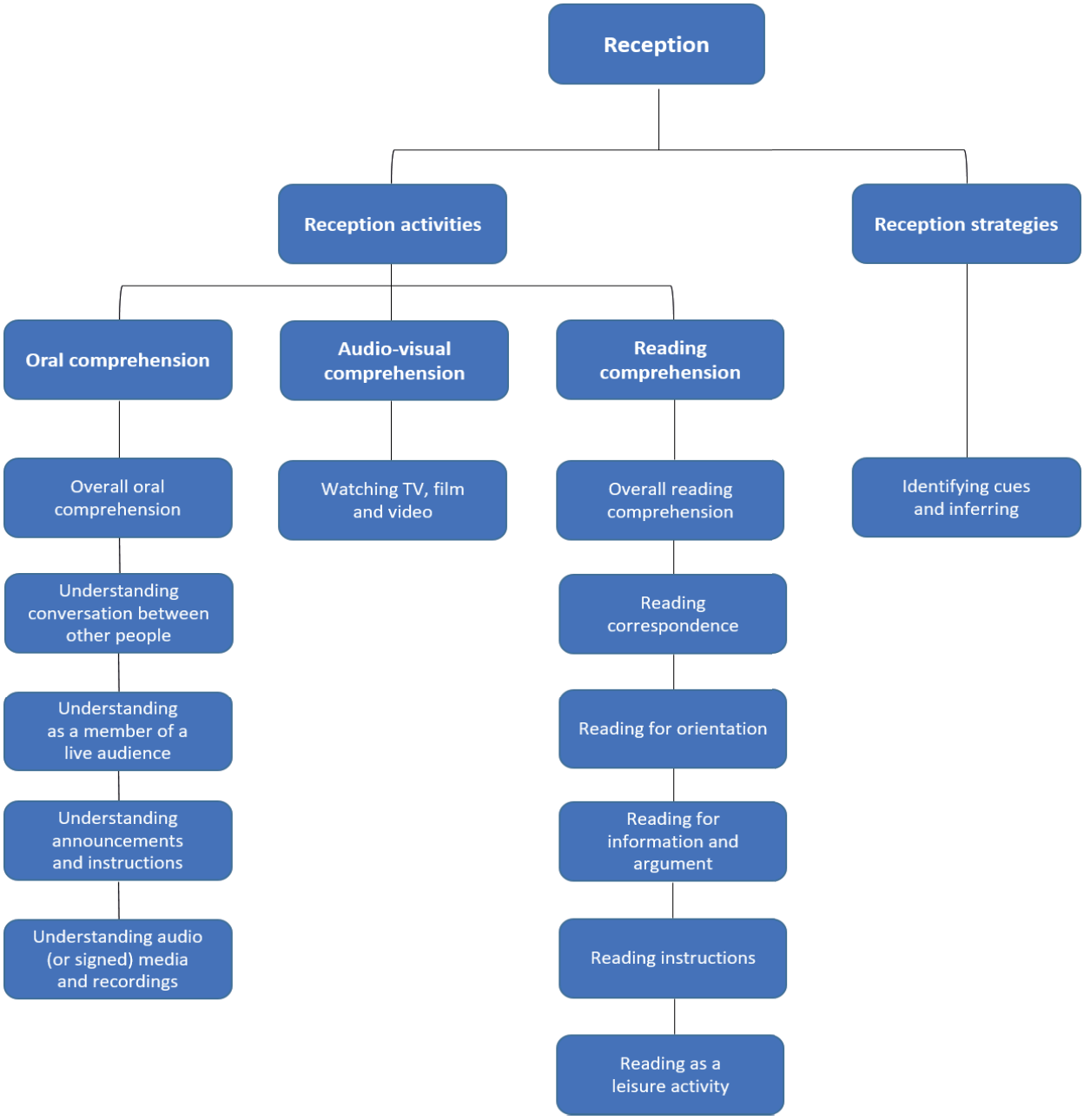
On their own, language competences do not constitute language use. It is the application of language competences by using language strategies in the context of communicative language activities that leads to communicative language use. The CEFR illustrative scales describe an extensive range of communicative language activities that language users typically engage with, and show the gradation of foreign language learner performance depending on CEFR proficiency level (Council of Europe, 2020, pp. 47-122[31]).

Communicative language activities include reception, production, interaction and mediation, as follows:

- **Reception** and **production** involve one-way communication where one main participant reads or listens silently, or delivers a speech or a piece of writing. In many educational or professional contexts, these types of receptive or productive language activities are common and play a fundamental role (e.g. learning and study purposes, oral presentations, written essays or reports, etc.). Reception activities and strategies are listed in Figure 2.2; see also Council of Europe (2020, pp. 47-60[31])). Production activities and strategies are listed in Figure 2.3; see also Council of Europe (2020, pp. 60-70[31]).[4]
- **Interaction** involves co-construction of meaning (or knowledge) where at least two participants alternate between reception and production as part of an oral or written exchange. Interaction is more than the sum of its constituent modes of communication because the interlocutors[5] are not simply speaking and listening to each other, or merely texting each other and reading each other's messages; the listener is also "already forecasting the remainder of the speaker's message and preparing a response" (Council of Europe, 2001, p. 14[33]) as does the reader of a text message in a two-way written chat. Figure 2.4 shows interaction activities and strategies; see also Council of Europe (2020, pp. 70-89[31]).
- **Mediation** is the act of facilitating communication between two or more individuals who are unable to communicate with each other directly. Typical examples involve translation and interpretation from one language to another (cross-linguistic mediation), but mediation may also occur within the same language. Mediation is categorised into three types (Council of Europe, 2020, pp. 90-117[31]):
    - **Mediating a (written, aural or visual) text**[6] is about passing on to someone else information about a text of any form to which they do not have access, often due to linguistic, cultural or other barriers. It can also be about mediating written, aural or visual material for oneself (e.g. taking notes during a class) or expressing opinions about texts. Activities here include relaying specific information; explaining data; processing text; translating a written text in speech; translating a written text in writing; note-taking (lectures, seminars, meetings, etc.); expressing a personal response to creative texts (including literature); and analysis and criticism of creative texts (including literature).
    - **Mediating concepts** involves making knowledge and concepts accessible to others who may be unable to access these directly on their own. It is a typical mode of communication in teaching, training, mentoring and parenting. Activities involve facilitating collaborative interaction with peers; collaborating to construct meaning; managing interaction; and encouraging conceptual talk.
    - **Mediating communication** refers to the process of facilitating understanding and shaping successful communication between language users who may have individual, sociocultural, sociolinguistic or intellectual differences in standpoint. It is concerned with diplomacy, negotiation and dispute resolution. Activities here comprise facilitating plurilingual space; acting as an intermediary in informal situations (with friends and colleagues); and facilitating communication in delicate situations and disagreements.

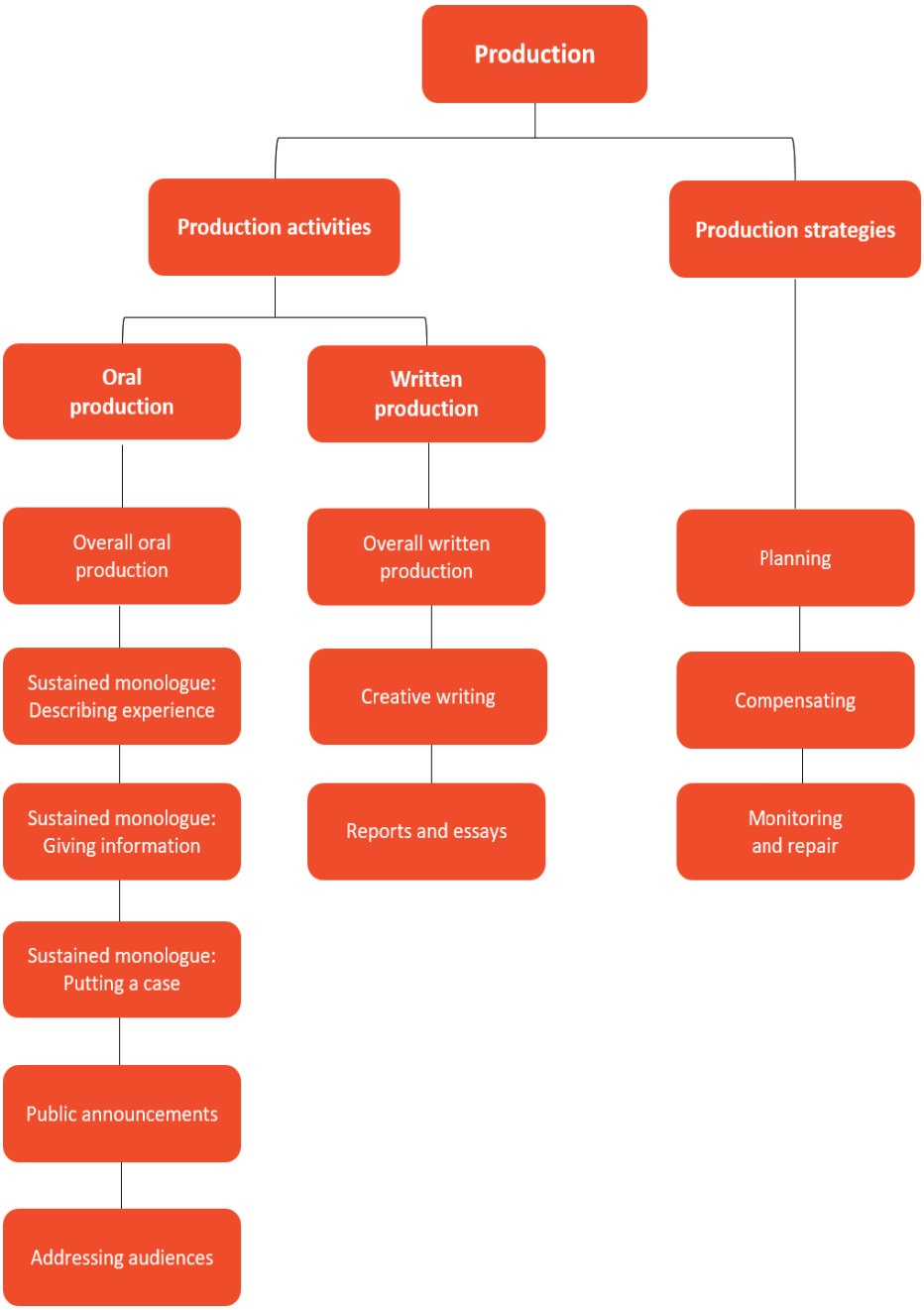
Mediation strategies relate to explaining a new concept – linking to previous knowledge, adapting language, breaking down complicated information – and to simplifying a text – amplifying a dense text and streamlining a text (Council of Europe, 2020, pp. 117-122[31]).

Figure 2.2. Communicative language activities and strategies: Reception



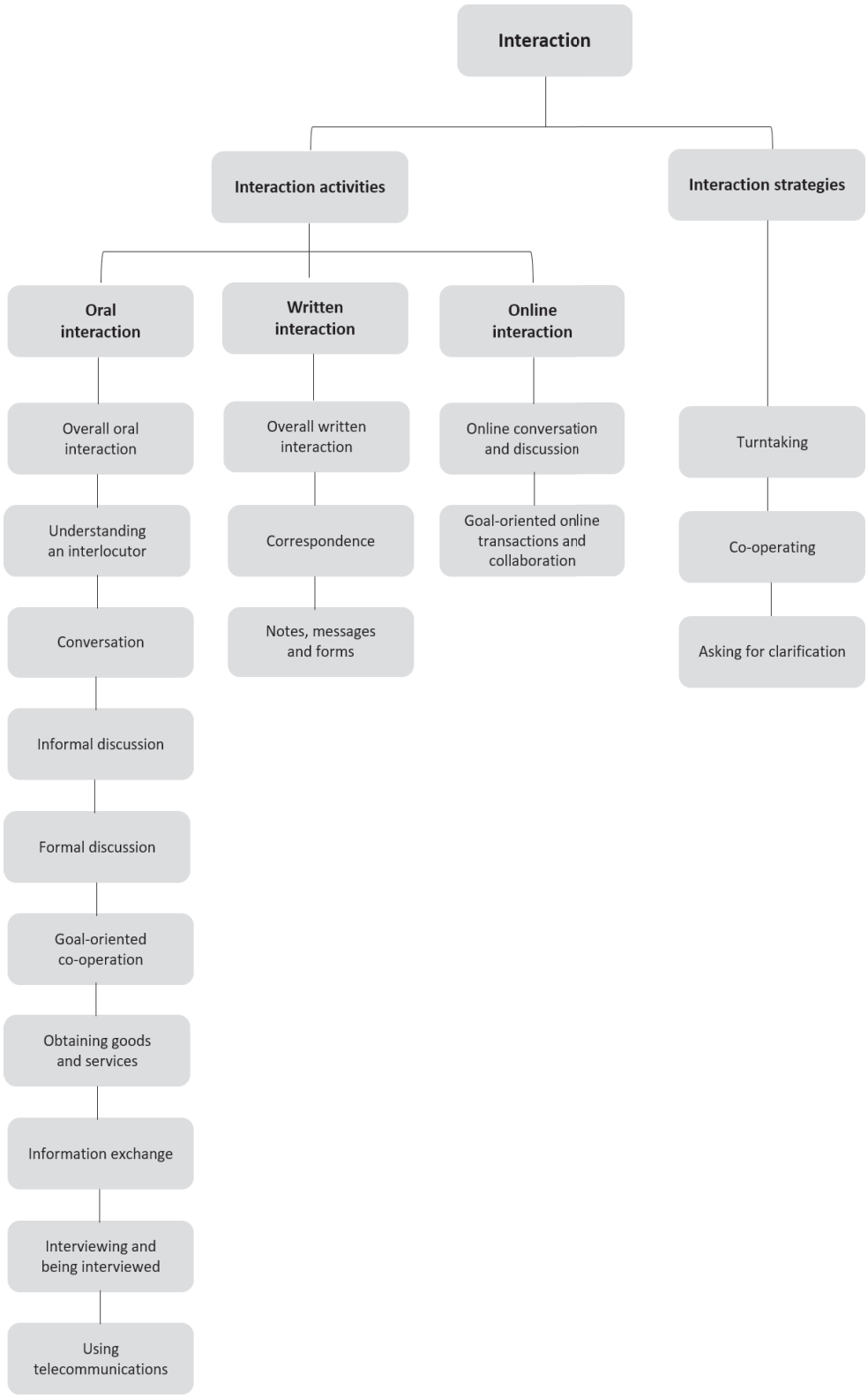
Source: Council of Europe (2020[31]), *Common European Framework of Reference for Languages: Learning, teaching, assessment - Companion Volume*, © Council of Europe, Strasbourg, p. 47.

Figure 2.3. Communicative language activities and strategies: Production



Source: Council of Europe (2020[31]), *Common European Framework of Reference for Languages: Learning, teaching, assessment - Companion Volume*, © Council of Europe, Strasbourg, p. 61.

Figure 2.4. Communicative language activities and strategies: Interaction



Source: Council of Europe (2020[31]), *Common European Framework of Reference for Languages: Learning, teaching, assessment - Companion Volume*, © Council of Europe, Strasbourg, p. 71.

Notes

1. In the FLA Framework, “task” is used to refer to any purposeful language activity that language users and learners engage with; this includes, in the context of the foreign language test, any assessment items or exercises (e.g. a set of questions about a text), which are also referred to as (assessment) tasks.

2. These four skills fall under what the CEFR describes as the communicative modes of reception, production, interaction and mediation. Reading and listening are involved in reception; speaking and writing in production; and all four skills can be involved in interaction and mediation.

3. This subdivision serves mainly descriptive purposes. As the CEFR states, these competences “are not separate ‘components’; [they] are always intertwined in any language use” (Council of Europe, 2020, p. 129[31]).

4. Traditionally, reading, listening, speaking and writing are referred to as the four language skills in the foreign language context. Here in Figures 2.2-2.4, they are listed under communicative language activities as the emphasis is on what foreign language learners use the language for as part of their foreign language proficiency. But, of course, they are both. Foreign language learners perform reading, listening, speaking, writing (and other language) activities by using the equivalent (reading, listening, speaking and writing) language skills.

5. Interlocutors are two or more persons who speak or converse with each other or text or write to each other.

6. As in the PISA Reading Literacy Framework, the term "text" in the PISA Foreign Language Assessment Framework is used to denote a written text. The only exception is in the context of mediation activities and strategies, where "text" refers to any form of text, including written, aural or visual material.

References

Ahmad, I. and M. Abidin (2020), "Developing an alternative listening comprehension test to benchmark Malaysian undergraduates' listening performances", *International Journal of Instruction*, Vol. 13/2, pp. 677-690, http://dx.doi.org/10.29333/iji.2020.13246a. [39]

Bannur, F., S. Abidin and A. Jamil (2015), "A Validation Process of ESP Testing Using Weir's Socio Cognitive Framework (2005)", *Procedia - Social and Behavioral Sciences*, Vol. 202, pp. 199-208, http://dx.doi.org/10.1016/j.sbspro.2015.08.223. [40]

Birdsong, D. (2006), "Age and Second Language Acquisition and Processing: A Selective Overview", *Language Learning*, Vol. 56/1, pp. 9-49. [5]

Boyd, E. and C. Taylor (2016), "Presenting Validity Evidence: The Case of the GESE", in Banerjee, J. and D. Tsagari (eds.), *Contemporary Second Language Assessment: Contemporary Applied Linguistics*, Bloomsbury Academic, London, http://dx.doi.org/10.5040/9781474295055.ch-002. [41]

Carlon, J. (2013), "The Implications of SLA Research for Latin Pedagogy: Modernizing Latin Instruction and Securing its Place in Curricula", *Teaching Classical Languages*, Vol. 4/2, pp. 106-122. [26]

Carpenter, D. (2000), "Reassessing the Goal of Latin Pedagogy", *The Classical Journal*, Vol. 95/4, pp. 391-395. [25]

Cenoz, J., F. Genesee and D. Gorter (2014), "Critical analysis of CLIL: Taking stock and looking forward", *Applied Linguistics*, Vol. 35/3, pp. 243-262, http://dx.doi.org/10.1093/applin/amt011. [15]

Council of Europe (2020), *Common European Framework of Reference for Languages: Learning, teaching, assessment - Companion Volume*, Council of Europe, Strasbourg, https://www.coe.int/en/web/education/-/common-european-framework-of-reference-for-languages-learning-teaching-assessment-companion-volume. [31]

Council of Europe (2018), *Common European Framework of Reference for Languages. Learning, Teaching, Assessment. Companion Volume with New Descriptors*, Council of Europe, Strasbourg, http://www.coe.int/lang-cefr. [49]

Council of Europe (2001), *Common European Framework of Reference for Languages. Learning, Teaching, Assessment*, Cambridge University Press, Cambridge, http://www.coe.int/lang-CEFR. [33]

Crystal, D. (2008), *A Dictionary of Linguistics and Phonetics*, Blackwell Publishing, Oxford. [23]

Dalton-Puffer, C., T. Nikula and U. Smit (2010), "Charting promises premises and research on content and language integrated learning", in Dalton-Puffer, C., T. Nikula and U. Smit (eds.), *Language Use and Language Learning in CLIL Classrooms*, John Benjamins Publishing. [13]

De Bot, K. and D. Larsen-Freeman (2011), "Researching Second Language Development from a Dynamic Systems Theory perspective", in Verspoor, M., K. De Bot and W. Lowie (eds.), *A Dynamic Approach to Second Language Development: Methods and techniques*, John Benjamins Publishing, https://doi.org/10.1075/lllt.29.01deb. [24]

De Houwer, A. (2009), *Bilingual First Language Acquisition*, Multilingualism Matters, Bristol. [2]

Dugdale, E. (2011), "Lingua Latina, Lingua Mea: Creative Composition in Beginning Latin", *Teaching Classical Languages*, Vol. 3/1, pp. 1-23. [28]

Dunlea, J. (2015), *Validating a set of Japanese EFL proficiency tests: demonstrating locally designed tests meet international standards*, University of Bedfordshire. [42]

ECML (n.d.), *Languages of Schooling*, https://www.ecml.at/Thematicareas/Languagesofschooling/tabid/2968/language/en-GB/Default.aspx (accessed on 17 June 2020). [30]

Ellis, R. (2015), *Understanding Second Language Acquisition*, Oxford University Press, Oxford. [18]

European Commission/EACEA/Eurydice (2017), *Key Data on Teaching Languages at School in Europe*, Publications Office of the European Union, Luxembourg, http://dx.doi.org/10.2797/04255. [32]

Eurydice (2006), *Content and Language Integrated Learning (CLIL) at School in Europe*, Eurydice, Brussels. [14]

Florescano, A. et al. (2011), "Developing affordable 'local' tests: the EXAVER project", in O'Sullivan, B. (ed.), *Language Testing: Theories and Practices*, Palgrave. [43]

Foster-Cohen, S. (2001), "First language acquisition... second language acquisition: 'What's Hecuba to him or he to Hecuba?'", *Second Language Research*, Vol. 17/4, pp. 329-344. [4]

Gass, S. and L. Selinker (2008), *Second Language Acquisition: An Introductory Course*, Routledge, New York. [20]

Harley, T. (2014), *The Psychology of Language, from Data to Theory*, Psychology Press, London. [1]

Ipek, H. (2009), "Comparing and Contrasting First and Second Language Acquisition: Implications for Language Teachers", *English Language Teaching*, Vol. 2/2, pp. 155-163. [11]

Khalifa, H. and C. Weir (2009), *Examining Reading: Research and Practice in Assessing Second Language Learning*, UCLES/Cambridge University Press, Cambridge. [36]

Koutropoulos, A. (2011), "Modernizing Classical Language Education: Communicative Language Teaching & Educational Technology Integration in Classical Greek", *Human Architecture: Journal of the Sociology of Self-Knowledge*, Vol. 9/3, pp. 55-69, http://scholarworks.umb.edu/humanarchitecturehttp://scholarworks.umb.edu/humanarchitecture/vol9/iss3/6. [27]

Lakshmanan, U. (2009), "Child Second Language Acquisition", in Ritchie, W. and B. Tej (eds.), *The New Handbook of Second Language Acquisition*, Emerald Group Publishing, Bingley. [3]

Lasagabaster, D. and J. Sierra (2010), "Immersion and CLIL in English: More differences than similarities", *English Language Teaching*, Vol. 64/4, pp. 367-375, http://dx.doi.org/10.1093/elt/ccp082. [12]

Lightbown, P. and N. Spada (2013), *How Languages are Learned*, Oxford University Press, Oxford. [19]

Littlewood, W. (1984), *Foreign and Second Language Learning: Language Acquisition Research and its Implications for the Classroom*, Cambridge University Press, Cambridge. [17]

Liu, J. and M. Pan (2019), "English Language Teaching in China: Developing Language Proficiency Frameworks", in Gao, X. (ed.), *Second Handbook of English Language Teaching*, Springer International Handbooks of Education, https://doi.org/10.1007/978-3-319-58542-0_28-1. [44]

Long, M. (1993), "Assessment Strategies for second language acquisition theories", *Applied Linguistics*, Vol. 14, pp. 225-249. [21]

Mehisto, P., M. Frigols and D. Marsh (2008), *Uncovering CLIL content and language integrated learning in bilingual and multilingual education*, Macmillan English. [22]

Moeller, A. and T. Catalano (2015), "Foreign Language Teaching and Learning", in Wright, J. (ed.), *International Encyclopedia for Social and Behavioural Sciences*, Pergamon Press, Oxford. [10]

Muñoz, C. and D. Singleton (2011), "A critical review of age-related research on L2 ultimate attainment", *Language Teaching*, Vol. 44/1, pp. 1-35, http://dx.doi.org/10.1017/S0261444810000327. [7]

Nikula, T., C. Dalton-Puffer and A. Llinares (2013), "CLIL classroom discourse: Research from Europe", *Journal of Immersion and Content-Based Language Education*, Vol. 1/1, pp. 70-100, http://dx.doi.org/10.1075/jicb.1.1.04nik. [16]

Oliveri, M. (2019), "Considerations for Designing Accessible Educational Scenario-Based Assessments for Multiple Populations: A Focus on Linguistic Complexity", *Frontiers in Education*, Vol. 4, http://dx.doi.org/10.3389/feduc.2019.00088. [45]

Patrick, R. (2019), "Comprehensible Input and Krashen's theory", *Journal of Classics Teaching*, Vol. 20/39, pp. 37-44, http://dx.doi.org/10.1017/s2058631019000060. [29]

Punchihetti, S. (2013), *First, second and foreign language learning: how distinctive are they from one another?*. [9]

Selvaruby, P., B. O'Sullivan and M. Watts (2008), "School-based assessment in Sri Lanka: ensuring valid processes for assessment-for-learning in physics", in Coll, R. and N. Taylor (eds.), *Education in Context: An International Perspective of the Influence of Context on Science Curriculum Development, Implementation and the Student-Experienced Curriculum*, Sense Publishers, Rotterdam. [46]

Shaw, S. and C. Weir (2007), *Examining Writing: Research and Practice in Assessing Second Language Writing*, UCLES/Cambridge University Press, Cambridge. [35]

Singleton, D. and L. Ryan (2004), *Language acquisition: The age factor*, Multilingual Matters. [6]

Su, L., C. Weir and J. Wu (2019), *English Language Proficiency Testing in Asia*, Routledge, New York, https://doi.org/10.4324/9781351254021. [47]

Taylor, L. (2011), *Examining Speaking: Research and Practice in Assessing Second Language Speaking*, UCLES/Cambridge University Press, Cambridge. [37]

Taylor, L. and A. Geranpayeh (2013), *Examining Listening: Research and Practice in Assessing Second Language Listening*, UCLES/Cambridge University Press, Cambridge. [38]

Tozlu, E. and A. Ünaldı (2018), "Striking the Balance between Validity and Reliability of a Listening Test in Turkish as a Second Language", *Boğaziçi Üniversitesi Eğitim Dergisi*, Vol. 34/1, pp. 49-73. [48]

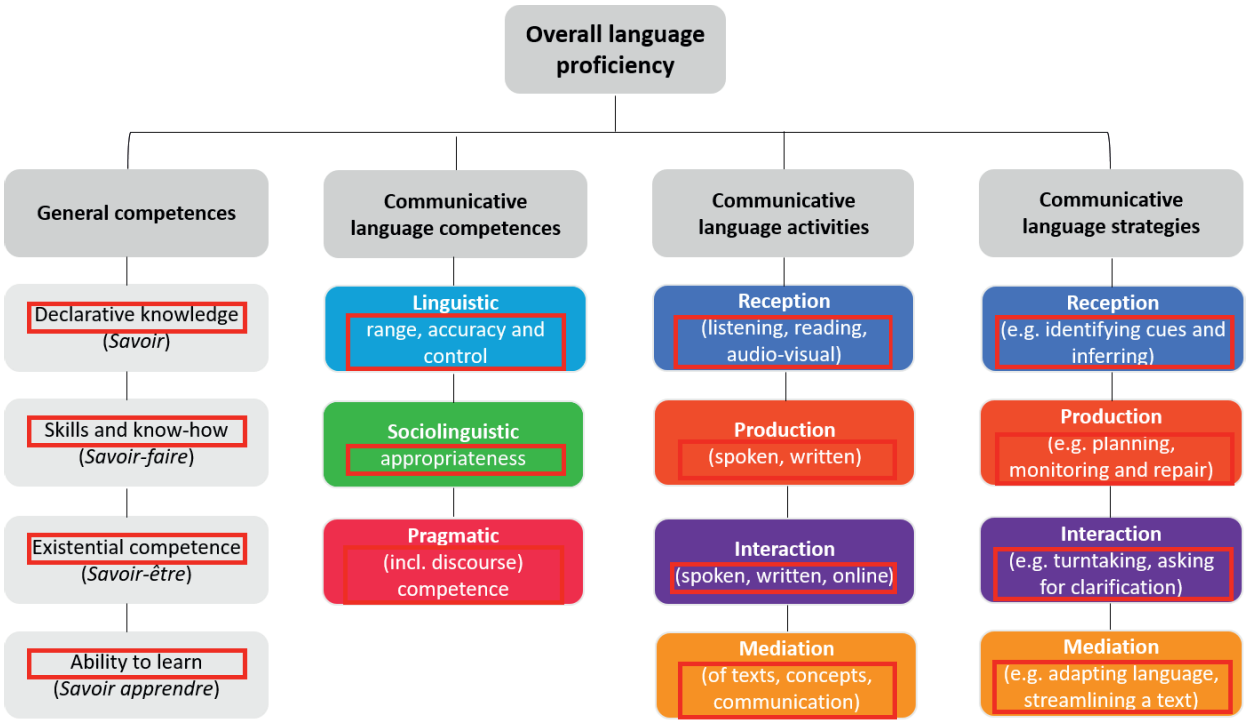
Weir, C. (2005), *Language Testing and Validation: An Evidence-Based Approach*, Palgrave McMillan, Hampshire. [34]

Yule, G. (2010), *The Study of Language*, Cambridge University Press, Cambridge. [8]

Annex 2.A. Adaptations to CEFR descriptive scheme

Enclosed in red boxes are the clarifications and examples that were added to the original CEFR descriptive scheme for language proficiency.

Annex Figure 2.A.1. Adaptations to CEFR descriptive scheme for language proficiency



Source: Adapted from Council of Europe (2020[31]), *Common European Framework of Reference for Languages: Learning, teaching, assessment - Companion Volume*, © Council of Europe, Strasbourg, p. 32.

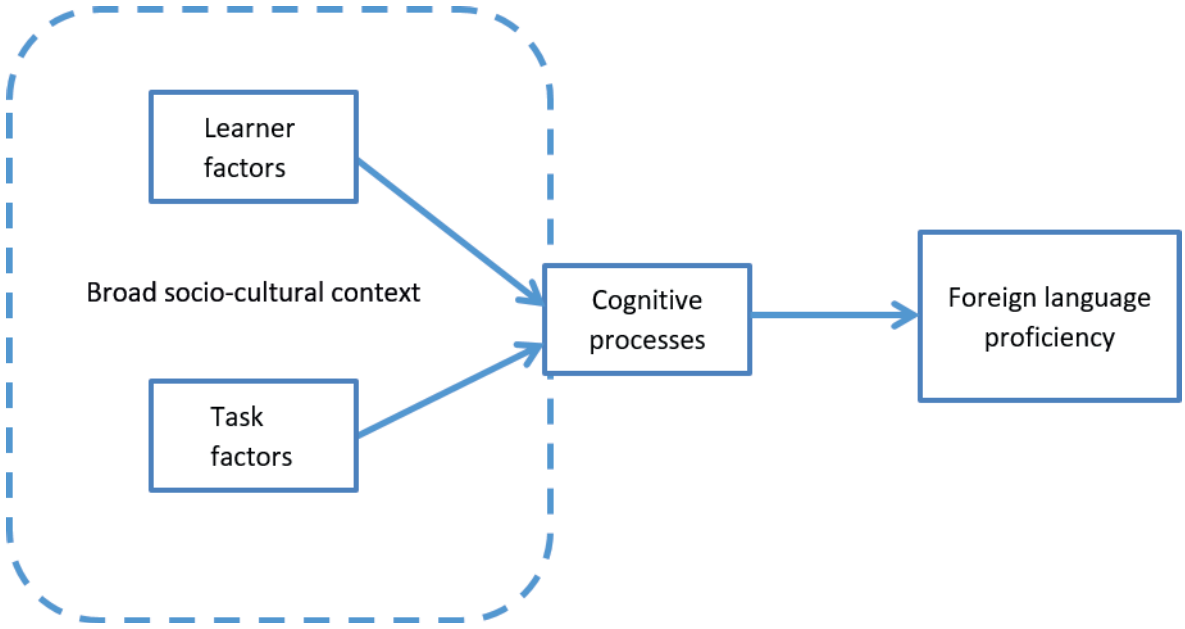
3 Organising the target constructs for the foreign language assessment

This chapter presents the theory underlying the reading, listening and speaking tests used to assess foreign language proficiency in the PISA Foreign Language Assessment. The assessment views foreign language proficiency as the joint outcome of two combined sources of influence that interact within a broad socio-cultural context: learner factors and the task or activity factors. These, in turn, determine the cognitive processes that take place when using a foreign language. The chapter describes the cognitive processes that learners employ when using a foreign language according to their proficiency level in the four skills of reading, listening, speaking and writing. It focuses on the first three skills, as writing will not be included in the PISA 2025 cycle.

Foreign language use and learning are highly complex, adaptive and diverse activities. In order to design an assessment that adequately represents the many facets of foreign language proficiency, the domain of foreign language proficiency is organised according to a set of dimensions. As in other Programme for International Student Assessment (PISA) frameworks, these dimensions will in turn determine the test design and, ultimately, the evidence about student proficiency that can be collected and reported. This PISA Foreign Language Assessment (FLA) Framework views foreign language proficiency as the joint outcome of two combined sources of influence: the learner and the task[1] or activity (which also includes texts for reading or audio input for listening). These, in turn, determine the cognitive processes that the learner employs when using a foreign language. Learner and task dimensions interact within a broad socio-cultural context, which can be thought of as the diverse range of situations in which foreign language learning and foreign language use occur.

Figure 3.1. illustrates these two main dimensions that together contribute to foreign language proficiency. A foreign language learner brings a number of learner factors (or characteristics) to language learning and use, which can include motivation, prior foreign language exposure and self-efficacy. The definition of learner factors also includes related dimensions that can affect learner behaviours and attainment, such as school environment, teachers' training, teaching and assessment practices (see Chapter 5 for the full list of learner factors the PISA FLA Framework will cover).

Figure 3.1. Factors that contribute to foreign language proficiency



Source: Authors' illustration

Proficiency in a foreign language is also a function of task factors (i.e. the reasons that motivate the foreign language learner to engage in language activity and the demands of language activity). Task factors include, but are not limited to, task purpose, function, complexity, time constraints and linguistic demands.[2] Based on their individual characteristics and their perception of task factors, foreign language learners apply a set of language-related and general cognitive processes in order to perform communicative language activities successfully. The language-related cognitive processes typically pertain to reading, listening, speaking and writing.

In designing the PISA Foreign Language Assessment, two points are important to consider. One is to ensure broad coverage of the language activities that foreign language learners typically engage with, both in and outside of school, and the other is to include tasks that elicit evidence about a range of CEFR (Common European Framework of Reference for Languages) levels. Broad coverage of the domain is ensured by including a wide variety of tasks and task purposes. Task difficulty is varied by manipulating task demands, which then dictate the deployment of a range of cognitive processes. Thus, by varying the dimensions of tasks used, the PISA Foreign Language Assessment aims to measure learners' foreign language proficiency through their ability to employ appropriate cognitive processes. While there may be individual differences in learner factors based on the experience and background of each foreign language learner, these are not manipulated in the Foreign Language Assessment but are captured through a background questionnaire (see Chapter 5).

The remainder of this chapter will describe the cognitive processes typically employed when individuals use a first language (L1) or a foreign language and will illustrate how these processes vary depending on a foreign language learner's proficiency level. The description will cover reading, listening, speaking and writing (with the intention of including further communication activities in future iterations of the framework). Learner factors that are captured in a separate questionnaire are discussed in Chapter 5. Task factors (including text factors) will be described in Chapter 4.

Reading

How do we read? A cognitive processing account of reading comprehension

Reading comprehension involves a complex set of processes, which are not necessarily employed in a straightforward nor linear fashion by readers, making it difficult to capture them in one model (Perfetti and Stafura, 2014[1]). There is a necessary trade-off between fully representing the dynamic nature of reading and providing a description of the reading processes that can be used to inform assessment practices and aid in describing proficiency. One such description is the Khalifa and Weir (2009[2]) reading model (Figure 3.2.), which synthesises work and models by authors in the field of cognitive psychology. These first language (L1) models can be (and have been) successfully adapted for second language (L2) learners.[3] This model provides a snapshot of core reading processes relevant to the PISA Foreign Language Assessment and is thus not exhaustive.

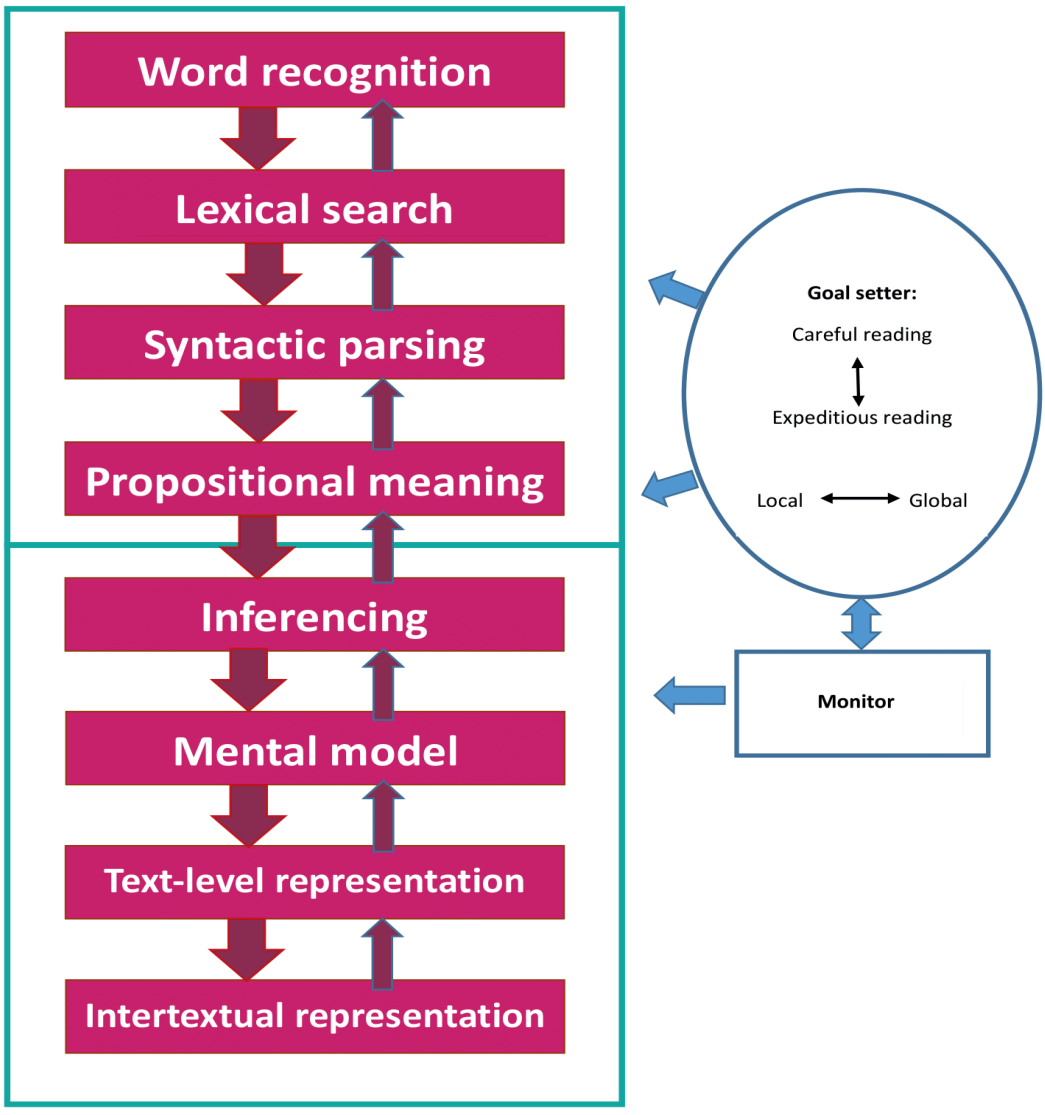
The model portrays reading as a set of metacognitive activities (e.g. deciding the purpose of reading) that trigger and monitor the cognitive processes needed to complete a reading task. Various forms of knowledge (e.g. lexical, syntactic, etc.) feed into the cognitive processes to enable successful reading comprehension.

The first step in reading comprehension is the metacognitive activity of goal setting (see the "Goal setter" oval in the right-hand column of the model), which determines the purpose of the reading activity and, subsequently, the type of reading to employ when faced with a text. As discussed in the PISA 2018 Reading Literacy Framework, readers read for a number of different purposes and to achieve different goals, whether it is to complete a school assignment, make an online purchase or simply read for pleasure, and they must adjust the way they read depending on the demands of the task (OECD, 2019[3]). The types of reading commonly employed fall along a continuum between careful and expeditious and can occur at the local or global level (Urquhart and Weir, 1998[4]). Speaking specifically:

- **Local comprehension** is the understanding of propositions at the sentence or clause level and is dependent on linguistic knowledge (i.e. lexical and grammatical ability) (Cohen and Upton, 2006[5]). It has also been referred to as comprehension of explicit information (Alderson, 2000[6]).
- **Global comprehension** is comprehension beyond the sentence level, such as understanding main ideas, and the links between propositions.

- **Careful reading** is slow, linear, incremental reading used to comprehend complete meanings. It can be employed either at the local level, i.e. within a sentence, or at the global level, i.e. beyond the sentence level up to the level of a complete text or across texts, where the purpose is to comprehend the main idea(s) and the overall text(s) or to determine how ideas relate to each other, and the author's aim.
- **Expeditious reading** is quick, selective reading used to access target information within a text or texts. Types of expeditious reading include skimming to obtain the gist or main overarching idea in a text (occurs at the global level); scanning to achieve specific reading goals, e.g. reading selectively to locate specific words in a text (occurs at the local level); and search reading to look for words in the same semantic field as the target information (can occur both at the local and global levels).

Figure 3.2. A model of reading comprehension



Source: Adapted from Khalifa, H. and C.J. Weir (2009[2]), *Examining Reading: Research and Practice in Assessing Second Language Learning*, UCLES/Cambridge University Press, Cambridge, p. 43.

In terms of the cognitive demands imposed by the different types of reading described above, there appears to be a cline from the local to the global level, and from careful to expeditious reading. This cline should also be reflected in terms of the reading texts and tasks foreign language readers are expected to successfully engage with across the CEFR levels. However, this does not preclude alternating between reading types in order to accomplish a reading purpose. For example, when reading expeditiously to locate specific information, the reader may then need to begin reading carefully at the local level to check that he/she has located the right information.

Once the goal setter has determined the purpose of reading, and, by extension, the type of reading to be deployed, this decision, in turn, determines the level(s) of cognitive processes activated during the reading activity (see the left-hand column of the model[4] in Figure 3.2.). Upon receiving visual input (the text), the reader engages, first, in word recognition, the matching of a written word form with a mental representation of the orthographic forms of the language. This is then followed by lexical access, the "retrieval of a lexical entry from the [mental] lexicon, containing stored information about a word's form [i.e. its orthography, phonology and possibly morphology] and meaning [i.e. its word class, syntactic structure and range of possible senses]" (Field, 2004, p. 151[7]). The reader then engages in syntactic parsing by organising words into phrases and larger units such as clauses and sentences, in order to understand the message of those units.

The next step is establishing propositional meaning at the clause or sentence level. Propositional meaning is a literal interpretation of what is written down. These first four levels are considered lower-level processes, as the reader is primarily focused on the linguistic code. However, establishing literal meaning may not be sufficient for full comprehension of a text, as links between ideas in a passage are often implicit. Therefore, the reader engages in higher-level processes, which centre on enriching meaning. If it is necessary to go beyond the literal meaning, the reader may need to engage in inferencing. It is not always possible to make every aspect of a piece of writing explicit, and writers make assumptions about what information can be reasonably expected to be shared with the reader and therefore does not need to be made explicit (Hyland, 2002[8]). Inferencing can take place at several different levels. Readers may need to infer the meaning of an ambiguous word, determine the antecedent of a pronoun through inference, recognise the implications of a phrase or add information "which is not stated in a text to impose coherence [and cohesion]" (Khalifa and Weir, 2009, p. 50[2]).

The next stage of processing, which often involves some element of inferencing, is building a mental model. According to Field (2004, p. 241[7]):

> *"[I]ncoming information has to be related to what has gone before, so as to ensure that it contributes to the developing representation of the text in a way that is consistent, meaningful and relevant. This process entails an ability to identify main ideas, to relate them to previous ideas, distinguishing between major and minor propositions, and to impose a hierarchical structure on the information in the text."*[5]

The final stages of processing involve creating a text-level representation for the text as a whole and, if more than one text is used, an intertextual representation:

> *"The skilled reader is able to recognise the hierarchical structure of the whole text, [to determine] which items of information are central to the meaning of the text, how the different parts of the text fit together and which parts … are important to the writer and reader purpose." (Weir and Khalifa, 2008, p. 7[9])*

In addition to the metacognitive activity of goal setting, which precedes and determines the cognitive reading processes being activated, another key component in reading comprehension is the metacognitive activity of monitoring (see the "Monitor" box in the right-hand column of the model in Figure 3.2.). With monitoring, readers check the effectiveness of their understanding at different points in the reading process (Sticht and James, 1984[10]). It is a complex mechanism that operates at various levels of the reading process. For instance, "[i]n decoding text, monitoring involves checking word recognition, lexical access, and syntactic parsing. Within meaning building, it can involve determining the success with which

the reader can extract the writer's intentions or the argument structure of the text" (Khalifa and Weir, 2009, p. 55[2]). Similarly, the PISA Reading Literacy Framework contends that skilled readers are able to monitor their progress towards accomplishing the aims or purposes of a reading activity and employ strategies or adjust reading processes in a "dynamic" way (OECD, 2019, p. 36[3]).

How do foreign language learners read? Cognitive processes across CEFR proficiency levels

Foreign language learners will employ different cognitive processes, depending on their proficiency levels. The Common European Framework of Reference for Languages (CEFR) provides insights into these processes through "can-do" descriptors[6]. Table 3.1 shows the CEFR overall reading comprehension scale that summarises reading ability at different proficiency levels (Council of Europe, 2020[11]).

Table 3.1. CEFR overall reading comprehension scale

CEFR level	Overall reading comprehension
C2	Can understand virtually all types of texts including abstract, structurally complex, or highly colloquial literary and non-literary writings. Can understand a wide range of long and complex texts, appreciating subtle distinctions of style and implicit as well as explicit meaning.
C1	Can understand in detail lengthy, complex texts, whether or not these relate to their own area of speciality, provided they can reread difficult sections. Can understand a wide variety of texts including literary writings, newspaper or magazine articles, and specialised academic or professional publications, provided there are opportunities for rereading and they have access to reference tools.
B2	Can read with a large degree of independence, adapting style and speed of reading to different texts and purposes, and using appropriate reference sources selectively. Has a broad active reading vocabulary, but may experience some difficulty with low-frequency idioms.
B1	Can read straightforward factual texts on subjects related to their field of interest with a satisfactory level of comprehension.
A2	Can understand short, simple texts on familiar matters of a concrete type which consist of high frequency everyday or job-related language.
	Can understand short, simple texts containing the highest frequency vocabulary, including a proportion of shared international vocabulary items.
A1	Can understand very short, simple texts a single phrase at a time, picking up familiar names, words and basic phrases and rereading as required.
Pre-A1	Can recognise familiar words/signs accompanied by pictures, such as a fast-food restaurant menu illustrated with photos or a picture book using familiar vocabulary.

Note: The CEFR makes a distinction between the "criterion levels" (e.g. A2) and the "plus levels" (e.g. A2+). "The latter are distinguished from the former by a horizontal line, as in [the A2 level in this table]. Plus levels represent a very strong competence at a level that does not yet reach the minimum standard for the following criterion level. Generally, features of the level above are starting to appear" (Council of Europe, 2018, p. 36[12]). There are three "plus levels" (A2+, B1+, B2+) and the CEFR scales cited throughout this chapter include these plus levels. Note, however, that this distinction is for descriptive purposes only, and that the PISA Foreign Language Assessment will not assess or report on the plus levels – see also Chapter 6 for more details on how the proficiency scales will be reported in PISA FLA.
Source: Council of Europe (2020[11]), *Common European Framework of Reference for Languages: Learning, teaching, assessment - Companion Volume*, © Council of Europe, Strasbourg, p. 54.

From Levels Pre-A1 to A2, the focus on short, simple texts of a concrete nature suggests that these learners are primarily able to cope with the lower-level reading processes (those up to establishing propositional meaning), with the main purpose employed being careful reading at the local level due to their limited grammatical and lexical knowledge. Learners at these reading levels are generally focused on word recognition, lexical access and syntactic parsing leading to the establishment of propositional meaning. Comprehension of a wider range of texts and increased level of comprehension is seen from Level B1, which suggests that readers at this level are able to engage in basic meaning construction, such

as making straightforward inferences and building a mental model, for example, to identify the author's purpose or identify attitudes or opinions. By Level B2, learners are able to read a wide range of texts independently by effectively deciding whether to read carefully or expeditiously either at the local or global level, depending on the type of text and the specific purpose for reading. The CEFR descriptors also suggest that learners at this level are able to engage in the higher-level processes to enhance propositional meaning. At Level C1, learners can understand the relationship between ideas in a text (i.e. building a mental model and textual representation). By Level C2, learners have the ability to create intertextual representations by comparing or contrasting multiple texts as well as "appreciating subtle distinctions in style" contained within these texts (Council of Europe, 2020, p. 54[11]). Finally, learners at this level are more able to recognise and use the organisation of the text as a whole to aid understanding and build a text-level structure.

Annex 3.A provides an overview of the cognitive processes which would need to be employed to achieve the key aspects of the CEFR descriptors for reading at each level.

Similarities and differences between the PISA Foreign Language Assessment Reading Framework and the PISA Reading Literacy Framework

This section briefly discusses the main differences and similarities between the reading model of the PISA FLA Framework and the PISA Reading Literacy Framework (OECD, 2019[3]). The PISA Foreign Language Assessment aims to assess proficiency in a language that is not the main language of instruction (for more details, see Chapter 2). In contrast, the PISA Reading Literacy Framework aims to assess reading proficiency in the main language of instruction. Throughout this chapter, L1 (first language) and language of instruction as assessed by the PISA Reading Literacy Framework will be used interchangeably.

Both frameworks describe the "what" and "how" of skilful reading; what a skilled reader is able to do and how he/she achieves this in either an L1 or a foreign language. The knowledge, skills, cognitive processes and strategies typically required for skilful reading are similar in L1 and at the highest levels of attainment in a foreign language. The cognitive processing model of reading comprehension discussed in the FLA Framework aims "to characterise the reading [processes and] behaviours available to the competent L1 reader which the foreign language reader might be expected to progressively approximate to as their proficiency level in a [foreign language] increases" (Weir and Khalifa, 2008, p. 5[9]).

Table 3.2. presents a broad mapping between the cognitive processes used in the PISA reading and foreign language frameworks, and the next two sections discuss their similarities and differences in greater detail.

Table 3.2. Broad mapping of reading cognitive processes of the PISA 2018 Reading Literacy Framework and the PISA 2025 Foreign Language Assessment Framework

Reading cognitive processes	
PISA 2018 Reading Literacy Framework	PISA 2025 Foreign Language Assessment Framework
Reading fluency: Reading words and text accurately and automatically Processing words and texts in order to comprehend the overall meaning of a text	Word recognition Lexical access Syntactic parsing
Locating information: Accessing and retrieving information within a text Searching for and selecting relevant text	(Expeditious reading)[1]

Reading cognitive processes	
PISA 2018 Reading Literacy Framework	PISA 2025 Foreign Language Assessment Framework
Understanding: Representing literal meaning Integrating and generating inferences	Establishing propositional meaning Inferencing Building a mental model Creating a text-level representation Creating an intertextual representation (Careful reading)
Evaluating and reflecting: Assessing quality and credibility Reflecting on content and form Detecting and handling conflict	*Not a main focus in the FLA framework*

1. Processes related to locating information are not explicitly assessed in the PISA FLA Framework but are involved, together with one or more of the remaining processes, in expeditious reading tasks.
Source: Authors.

Similarities and differences between the cognitive processes employed by first language and foreign language readers

How the CEFR conceptualises what foreign language readers can do at each proficiency level has similarities with how reading literacy is described in the PISA Reading Literacy Framework (OECD, 2019, p. 55[3]). For example, the lowest level in the PISA 2018 reading proficiency scale states:

> *"Readers at Level 1b can locate a single piece of explicitly stated information in a prominent position in a short, syntactically simple text with a familiar context and text type, such as a narrative or a simple list. Texts in Level 1b tasks typically provide support to the reader, such as repetition of information, pictures or familiar symbols. There is minimal competing information. Level 1b readers can interpret texts by making simple connections between adjacent pieces of information." (OECD, 2019, p. 55[3])*

There are parallels between this level and the CEFR Level A1 descriptor. However, there are important differences to keep in mind. A 15-year-old foreign language learner who achieves a Level A1 in reading may not necessarily have difficulties with the cognitive processes needed to read effectively. Rather they may not have sufficient lexical and/or grammatical knowledge in the foreign language to demonstrate higher-level cognitive processes. It can be challenging to determine in a foreign language whether a learner has language difficulties or reading difficulties because of the interdependence between linguistic competence and the higher-level reading processes involved in deriving meaning from a text, and because of the inter-relationship between first language and foreign language reading competence (Alderson, 1984[13]; Cummins, 1979[14]). The view that general cognitive processes involved in reading are universal and transferrable from one language to another may imply that learners do not necessarily need to learn how to read in the foreign language because they can transfer their first language reading skills to the foreign language (Cummins, 1991[15]). However, it has also been found that first language cognitive abilities cannot be transferred until a proficiency threshold is reached in the foreign language (Cummins, 1976[16]). PISA may represent an opportunity to assess the reading proficiency of foreign language learners both in their first language (based on the PISA reading test, administered in the main language of instruction) and in a foreign language (based on the PISA FLA reading test).

This perspective on language transfer of cognitive processes in reading is, however, further complicated by language-specific variations in writing systems, which have been shown to influence reading processes in the first language (Perfetti and Harris, 2013[17]). When children learn to read, they acquire cognitive reading processes that are considered universal but also acquire text-processing procedures that are specific to their first language (Koda, 2005[18]). This can pose challenges for learners if the first language and the foreign language do not share the same writing system or orthographies. For example, research

has shown that word recognition is slower and less accurate when writing systems are different, which may affect meaning construction (Koda, 1989[19]).

Similarities and differences between the reading literacy models

Both the PISA 2018 Reading Literacy Framework and the PISA 2025 Foreign Language Assessment Framework highlight a number of similar cognitive processes required for reading. For example, reading fluency in the PISA Reading Literacy Framework is a concept that relies on the effectiveness of decoding, word recognition, lexical access and syntactic parsing, which are fundamental to foreign language reading, and are the lower-level cognitive processes found in the foreign language reading model. Naturally, these processes have more prominence in the foreign language framework because they are the ones foreign language learners typically employ at the lower levels of proficiency, and can require more effort and are less automatic at the initial stages of reading in a foreign language than in one's main language of instruction. Reading fluency in a foreign language, like in the first language, is about speed and automaticity. The higher-level cognitive processes of the foreign language framework correspond more to the process of understanding information in the PISA 2018 Reading Literacy Framework. This is where learners identify literal meaning (i.e. establishing propositional meaning), enrich meaning through inferencing and if required to do so, build a mental model of the text, create a text-level structure, and integrate information from more than one text to create an intertextual representation.

Where the two models diverge relates to the emphasis on evaluating and reflecting on information in the PISA Reading Literacy Framework and on the goal setter in the reading model in the PISA FLA Framework. In the former framework, evaluating and reflecting on information represents an important reading activity that centres on critical reflection on a text in terms of its content, quality and form. These processes have less prominence in the PISA FLA Framework, as the ability to read critically and reflectively in the foreign language also depends on a certain threshold of foreign language knowledge and proficiency. Such processes are captured only at the C levels in the CEFR, which limits their inclusion in the PISA Foreign Language Assessment.

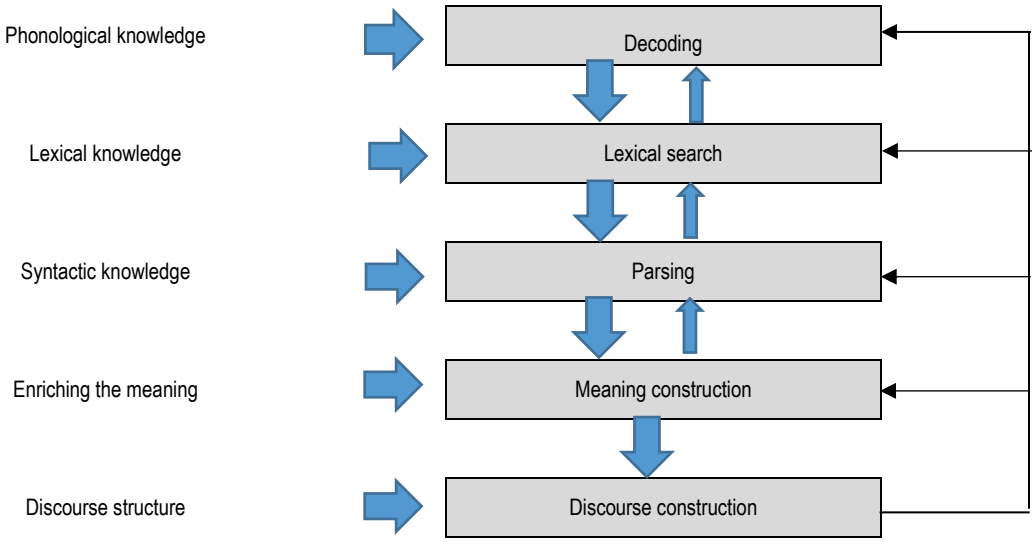
Finally, although the purpose for reading and types of reading processes can be found in both frameworks, they are represented differently. Both frameworks consider the identification of reading goals as a form of metacognition outside the core cognitive processes. However, in the PISA Reading Literacy Framework, types of reading are embedded within locating information (i.e. expeditious reading) and understanding information (i.e. careful reading). The PISA FLA Framework, on the other hand, makes this metacognitive decision more explicit through the inclusion of the goal setter, which is where the reader decides on the goals for reading and what type of reading to engage in. By doing so, the PISA FLA Framework can account for cases where learners may engage in an ineffective type of reading because they lack control over the lower-level reading processes.

Listening

How do we listen? A cognitive processing account of listening comprehension

The nature of listening fundamentally involves deriving meaning from acoustic input. The variability of input and the need to process it in real time contribute to the complexity of the skill. Field (2013[20]) has synthesised the listening research and proposes a model to describe the listening process. His model draws on and builds upon Cutler and Clifton's L1 listening model (1999[21]) and Field's L2 listening model (2009[22]). There are five major levels of processing in this model (see Figure 3.3.).

Figure 3.3. A model of listening comprehension



Note: This model is a simplified version of the listening model presented by Field (2013[20]).
Source: Adapted from Field, J. (2013[20]), "Cognitive validity", in Taylor, L. and A. Geranpayeh (eds.), *Examining Listening: Research and Practice in Assessing Second Language Listening*, UCLES/Cambridge University Press, Cambridge.

The first level is decoding, which involves converting the incoming acoustic sounds into phonemes and syllables using one's phonological knowledge of the target language. Decoding includes identifying which syllables carry stress, as this feeds into the next level, lexical search. As listeners decode the input, they begin to search their mental lexicon for the best word match. Because word boundaries are not clearly marked in speech, listeners use a number of strategies to narrow down the most likely word being spoken. They use stress and intonation to identify segment boundaries, and word frequency and lexical associations to help narrow down and determine the most likely words being spoken. Listeners also begin to consider what sense or meaning of the words is being used. The next level is parsing, which involves imposing a syntactic structure onto the words to identify the literal meaning of an utterance.

However, the literal meaning often does not accurately reflect the intended meaning. A feature of spoken language is that quite a lot of information is left unstated because there is an assumption by speakers that a great deal of knowledge is shared with the listener (Schank and Abelson, 1977[23]). Therefore, listeners need to determine the relevance of an utterance or its intended meaning, which requires higher-level cognitive processing and is referred to as meaning construction. Listeners use pragmatic, external and topic knowledge to interpret the speaker's intentions, infer information that has not been included, and use discourse features to make connections between utterances. Pragmatic knowledge helps the listener understand the purpose of the utterance beyond the syntactic form. For example, the utterance "I am cold" could be a request to turn up the heat, a statement of fact or an invitation to start a conversation depending on the context of the utterance. Listeners also use their knowledge of the speaker, the world and the situation to further interpret the message. They use their topic knowledge to make inferences about information that is not stated and interpret the antecedents of reference markers.

Finally, at the last level, discourse construction, the listener makes choices about the relevance of all the information received to decide on the overall main message and put the utterances into a larger discourse structure, which is used to aid recall. The features of discourse construction have been drawn from reading research (Gernsbacher, 1990[24]; Oakhill and Garnham, 1988[25]; Van Dijk and Kintsch, 1983[26]), and include deciding on the relevance of new information (i.e. selection), making links between new and

previously stated information (i.e. integration), self-monitoring to check that the interpretations being made are accurate as new information is presented, and building a hierarchical structure to separate the main points from subsidiary points.

In this listening model, decoding, lexical search and parsing are considered lower-level processes, and meaning and discourse construction are higher-level processes. The arrows in the model highlight that listening is an "online" activity in which the listener is creating hypotheses in real time as to what is being said and keeping them in mind until they are confirmed or replaced with new hypotheses. Therefore, listeners can process information at more than one level at a time, so it is not a linear process.

How do foreign language learners listen? Cognitive processes across CEFR proficiency levels

Depending on the language level of the listener, he/she will face different challenges when listening. The CEFR contains a number of illustrative listening scales that provide guidance on the role of the listener (e.g. participant, eavesdropper), what a learner is expected to understand and the characteristics of the audio texts. The CEFR overall listening comprehension descriptors in Table 3.3 (Council of Europe, 2020, p. 48[11]) provide a summary of the different scales and indirectly provide some indications of the cognitive processes learners can cope with at each level.[7]

Table 3.3. CEFR overall listening comprehension scale

CEFR level	Overall listening comprehension
C2	Can understand with ease virtually any kind of language, whether live or broadcast, delivered at fast natural speed.
C1	Can understand enough to follow extended discourse on abstract and complex topics beyond their own field, though they may need to confirm occasional details, especially if the variety is unfamiliar. Can recognise a wide range of idiomatic expressions and colloquialisms, appreciating register shifts. Can follow extended discourse even when it is not clearly structured and when relationships are only implied and not signalled explicitly.
B2	Can understand standard language or a familiar variety, live or broadcast, on both familiar and unfamiliar topics normally encountered in personal, social, academic or vocational life. Only extreme [auditory/visual] background noise, inadequate discourse structure and/or idiomatic usage influence the ability to understand.
	Can understand the main ideas of propositionally and linguistically complex discourse on both concrete and abstract topics delivered in standard language or a familiar variety, including technical discussions in their field of specialisation. Can follow extended discourse and complex lines of argument, provided the topic is reasonably familiar, and the direction of the argument is signposted by explicit markers.
B1	Can understand straightforward factual information about common everyday or job-related topics, identifying both general messages and specific details, provided people articulate clearly in a generally familiar variety.
	Can understand the main points made in clear standard language or a familiar variety on familiar matters regularly encountered at work, school, leisure, etc., including short narratives.
A2	Can understand enough to be able to meet needs of a concrete type, provided people articulate clearly and slowly.
	Can understand phrases and expressions related to areas of most immediate priority (e.g. very basic personal and family information, shopping, local geography, employment), provided people articulate clearly and slowly.
A1	Can follow language which is very slow and carefully articulated, with long pauses for them to assimilate meaning. Can recognise concrete information (e.g. places and times) on familiar topics encountered in everyday life, provided it is delivered slowly and clearly.
Pre-A1	Can understand short, very simple questions and statements, provided they are delivered slowly and clearly and accompanied by visuals or manual gestures to support understanding and repeated if necessary. Can recognise everyday, familiar words/signs, provided they are delivered clearly and slowly in a clearly defined, familiar everyday context. Can recognise numbers, prices,

	dates and days of the week, provided they are delivered slowly and clearly in a defined, familiar everyday context.

Note: The 2020 CEFR uses the term "oral comprehension" for this scale, but to simplify this framework and in accordance with the terminology used throughout, the term "listening comprehension" has been used instead.
Source: Council of Europe (2020[11]), *Common European Framework of Reference for Languages: Learning, teaching, assessment - Companion volume*, © Council of Europe, Strasbourg, p. 48.

The CEFR descriptors suggest that learners at Levels Pre-A1 to A2 are primarily able to cope with the lower-level cognitive processes in the listening model: decoding, lexical search and parsing because of the focus on understanding concrete, factual information that is topically familiar. They rely on modified input in terms of speech rate, vocabulary choice and the grammatical forms included to aid in comprehension, which suggests learners at these levels do not have sufficient automaticity of the lower-level processes to engage in the higher-level processes involved in enriching meaning. By Level B2, learners are showing signs that they can start to enrich the meaning of an utterance to arrive at the intended meaning and engage in discourse construction because they can understand main ideas, even on abstract topics, and follow lines of argument. The cultural knowledge that can be understood and used to more effectively enrich meaning and build a discourse representation expands markedly by Level B2.

Annex 3.A provides an overview of the cognitive processes which would need to be employed to achieve the key aspects of the CEFR descriptors for listening at each level.

Speaking

How do we speak? A cognitive processing account of speaking

The process of speaking, of assembling speech, is characterised by both physiological constraints, such as articulating speech at speed, and psychological constraints, such as the processing capacity (short-term memory and planning time) available to manage incoming and outgoing messages and utterances, while simultaneously attending to different aspects of performance. To describe the cognitive processes involved in speaking, the PISA FLA Framework will use Field's (2011[27]) model, which offers a succinct account of speaking processes based on Levelt's (1989[28]) comprehensive cognitive model of speech. Field's model (Figure 3.4.) depicts five main levels of speech processing (in the right-hand column) and the range of information sources that feed into each level.

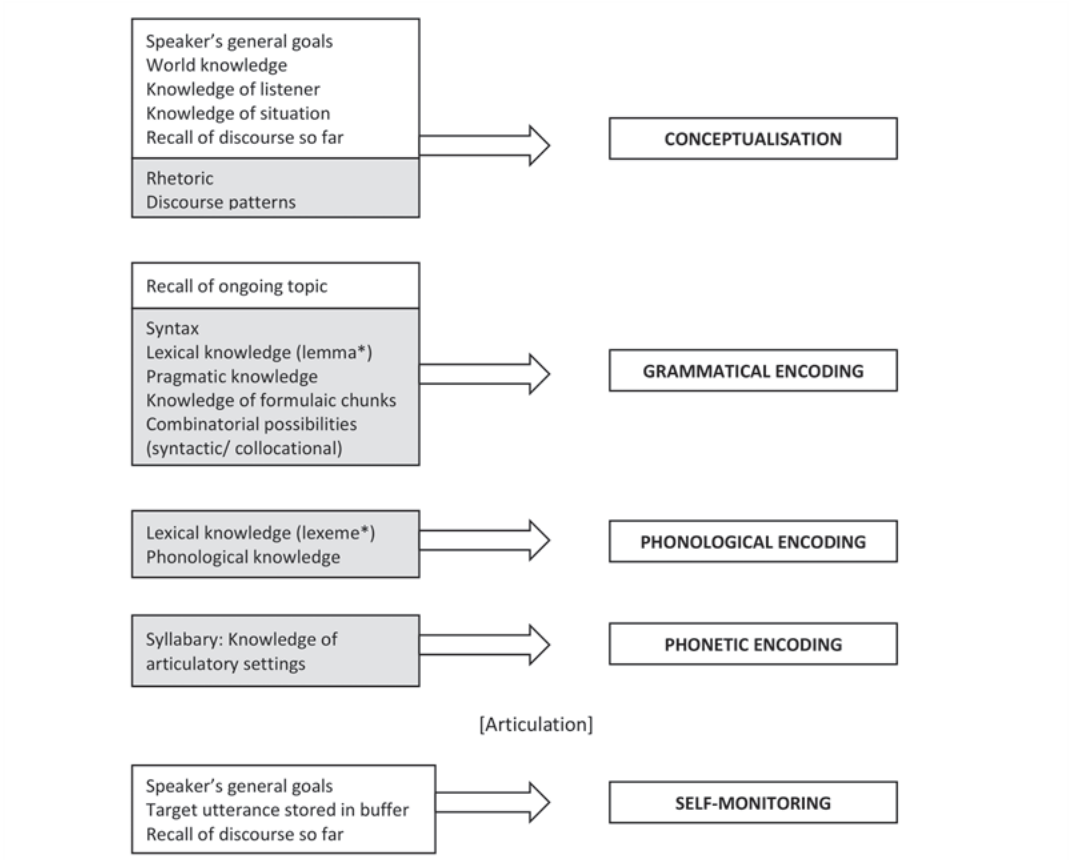
Speech starts with **conceptualisation**, which in itself entails macro- and micro-planning. Macro-planning occurs when the speaker anticipates a set of speech acts and is constrained by the limitations of planning time and working memory. Micro-planning occurs locally and is concerned with the form of the next utterance, which may include shared and unshared information, and accommodation to point of view, as well as the place of this next utterance in shared discourse.

The next step is **grammatical encoding**, which entails constructing abstract syntactic patterns and locating the necessary lexical items. This is followed by **phonological encoding**, the conversion of these abstract patterns and lexical items to phonological form with the retrieval of appropriate phonology from memory. Once the phonological form has been retrieved, **phonetic encoding and articulation** commence. Phonetic encoding involves adjusting the retrieved phonological sequence to make articulation easier (e.g. by introducing assimilation[8]), converting each of the syllables to a set of neural instructions for the articulators, and storing these instructions in a buffer while the clause is being articulated. Articulation is the actual production of an utterance.

Finally, **self-monitoring** takes place. It refers to the speaker's assessment of how effectively their plans for each utterance were achieved during its assembly. It compares the rhetorical impact of what was said

to the goals of the speaker at the conceptualisation stage. Skilled speakers check that their utterance is accurate, clear and has fulfilled their pragmatic intentions.

Figure 3.4. A model of speech production



Note: *In the 1999 update of his model, Levelt distinguishes between three components in a lexical entry in the mind storing information about a word. There is a semantic component, which enables a match to be made between a meaning and the target word: a lemma containing syntactic information about the word (its word class and combinatorial possibilities); and a lexeme containing information about the word's phonological form and morphology.
Source: Field, J. (2011[27]), "Cognitive validity", in Taylor, L. (ed.), *Examining Speaking: Research and Practice in Assessing Second Language Speaking*, UCLES/Cambridge University Press, Cambridge.

How do foreign language learners speak? Cognitive processes across CEFR proficiency levels

How do the speaking cognitive processes mentioned above vary for the foreign language speaker depending on their CEFR proficiency level? The "Overall spoken production" and "Overall interaction production" scales (Table 3.4 and Table 3.5) summarise the main features captured in the different CEFR speaking scales, and they are used as a jumping-off point to explore the cognitive demands on the foreign language speaker. These increase gradually from the basic tasks of Level Pre-A1, where the foreign language speaker is expected to ask and answer short, predicted questions about themselves and very familiar daily matters, to the highly sophisticated cognitive demands of Level C2, where it is anticipated that the foreign language speaker is able to produce smoothly flowing well-structured speech that conveys finer shades of meaning with a wide range of linguistic and discourse means. Although all or most of the five speaking cognitive processes outlined in the Field (2011[27]) model are required for speech production

at all proficiency levels, the task demands differ substantially between the levels and the quality of the output produced by learners. These differences are described below.

Table 3.4. CEFR overall spoken production scale

CEFR level	Overall spoken production
C2	Can produce clear, smoothly flowing, well-structured discourse with an effective logical structure which helps the recipient to notice and remember significant points.
C1	Can give clear, detailed descriptions and presentations on complex subjects, integrating sub-themes, developing particular points and rounding off with an appropriate conclusion.
B2	Can give clear, systematically developed descriptions and presentations, with appropriate highlighting of significant points, and relevant supporting detail.
	Can give clear, detailed descriptions and presentations on a wide range of subjects related to their field of interest, expanding and supporting ideas with subsidiary points and relevant examples.
B1	Can reasonably fluently sustain a straightforward description of one of a variety of subjects within their field of interest, presenting it as a linear sequence of points.
A2	Can give a simple description or presentation of people, living or working conditions, daily routines, likes/ dislikes, etc. as a short series of simple phrases and sentences linked into a list.
A1	Can produce simple, mainly isolated phrases about people and places.
Pre-A1	Can produce short phrases about themselves, giving basic personal information (e.g. name, address, family, nationality).

Note: The 2020 CEFR uses the term "oral production" for this scale, but to simplify this framework and in accordance with the terminology used throughout, the term "spoken production" has been used instead.
Source: Council of Europe (2020[11]), *Common European Framework of Reference for Languages: Learning, teaching, assessment - Companion volume*, © Council of Europe, Strasbourg, p. 62.

Table 3.5. CEFR overall spoken interaction scale

CEFR level	Overall spoken interaction
C2	Has a good command of idiomatic expressions and colloquialisms with awareness of connotative levels of meaning. Can convey finer shades of meaning precisely by using, with reasonable accuracy, a wide range of modification devices. Can backtrack and restructure around a difficulty so smoothly that the interlocutor is hardly aware of it.
C1	Can express themselves fluently and spontaneously, almost effortlessly. Has a good command of a broad lexical repertoire allowing gaps to be readily overcome with circumlocutions. There is little obvious searching for expressions or avoidance strategies; only a conceptually difficult subject can hinder a natural, smooth flow of language.
B2	Can use the language fluently, accurately and effectively on a wide range of general, academic, vocational or leisure topics, marking clearly the relationships between ideas. Can communicate spontaneously with good grammatical control without much sign of having to restrict what they want to say, adopting a level of formality appropriate to the circumstances.
	Can interact with a degree of fluency and spontaneity that makes regular interaction, and sustained relationships with users of the target language, quite possible without imposing strain on either party. Can highlight the personal significance of events and experiences, and account for and sustain views clearly by providing relevant explanations and arguments.
B1	Can communicate with some confidence on familiar routine and non-routine matters related to their interests and professional field. Can exchange, check and confirm information, deal with less routine situations and explain why something is a problem. Can express thoughts on more abstract, cultural topics such as films, books, music, etc.
	Can exploit a wide range of simple language to deal with most situations likely to arise while travelling. Can enter unprepared into conversation on familiar topics, and express personal opinions and exchange information on topics that are familiar, of personal interest or pertinent to everyday life (e.g. family, hobbies, work, travel and current events).
A2	Can interact with reasonable ease in structured situations and short conversations, provided the other person helps if necessary. Can manage simple, routine exchanges without undue effort;

CEFR level	Overall spoken interaction
	can ask and answer questions and exchange ideas and information on familiar topics in predictable everyday situations.
	Can communicate in simple and routine tasks requiring a simple and direct exchange of information on familiar and routine matters to do with work and free time. Can handle very short social exchanges but is rarely able to understand enough to keep conversation going of their own accord.
A1	Can interact in a simple way but communication is totally dependent on repetition at a slower rate, rephrasing and repair. Can ask and answer simple questions, initiate and respond to simple statements in areas of immediate need or on very familiar topics.
Pre-A1	Can ask and answer questions about themselves and daily routines, using short, formulaic expressions and relying on gestures to reinforce the information.

Note: The 2020 CEFR uses the term "oral interaction" for this scale, but to simplify this framework and in accordance with the terminology used throughout, the term "spoken interaction" has been used instead.
Source: Council of Europe (2020[11]), *Common European Framework of Reference for Languages: Learning, teaching, assessment - Companion volume*, © Council of Europe, Strasbourg, p. 72.

At the initial stage of **conceptualisation**, both macro-planning (i.e. the retrieval of information and generation of ideas) and micro-planning (e.g. integration of ideas and utterances into a discourse framework) place considerable cognitive processing demands on the foreign language speaker (Field, 2011[27]). The lower the proficiency level of the foreign language speaker, the higher these demands are. As a result, the expectations for the provision of ideas and information on the part of the foreign language speaker differ considerably across the CEFR levels (Field, 2011, p. 88[27]). In the early levels, there is an emphasis on personal and everyday information, which is conceptually simple and easily accessed. At higher CEFR levels, such as C1 and C2, task content is broadened to include unfamiliar topics, and learners are required to engage in more abstract discussion. Moreover, as foreign language speakers go up the CEFR levels, they are expected to demonstrate an increasing ability to relate ideas and utterances to the wider discourse and to have an increasing awareness of information shared by the interlocutor.

At the stage of **syntactic encoding**, the cognitive demands placed on the foreign language speaker do not derive only from the syntactic and lexical complexity of the utterance(s) to be produced, but also from the ease with which the foreign language speaker can map the function[9] they want to perform to the linguistic pattern that best expresses that function (Field, 2011[27]). Functions that are familiar, frequent and concrete characterise levels up to B1, and it is at Level B1 when learners can begin to engage in interactional functions such as agreeing or disagreeing. From Level B2 upwards, the number and complexity of functions rise with proficiency level, most noticeably with the ability to deal with modality. However, as Galaczi and ffrench (2011, p. 165[29]) indicate, the same functions can be appropriate across levels but the "degree of accuracy and complexity with which [the learners] can express their views" differs as proficiency increases. Table 3.6 provides an overview of key functions by CEFR proficiency level.

Table 3.6. Overview of key speaking functions by CEFR proficiency level

CEFR level	Overview	Selected examples[1]
C2	Learners at this level have a deeper appreciation of the full range of functional exponents. They can demonstrate sensitivity to their audience and are stylistically proficient.	Evaluating
C1	Learners at this level can meet the functional requirements almost effortlessly. They should have an increasing awareness of how to express functions in stylistically different ways.	Expressing and justifying opinions Hypothesising Summarising
B2	Learners at this level are expected to deal with functions with a greater degree of range, accuracy and spontaneity.	Comparing and contrasting

		Giving an opinion Negotiation Speculating Decision making Expressing modality
B1	A slightly more demanding understanding of functions is required in terms of range, accuracy and the ability to manage interactions.	Referring to past, present or future Suggesting Agreeing and disagreeing Giving cause, reason or example
A2	Basic functions relating to personal information, everyday activities and simple social interaction.	Presenting facts Requesting facts

1. Learners at each level should be able to cope with all the functions in the previous level.
Source: Field, J. (Field, 2011[27]), "Cognitive validity", in Taylor, L. (ed.), *Examining Speaking: Research and Practice in Assessing Second Language Speaking*, UCLES/Cambridge University Press, Cambridge, p. 92; Galaczi, E. D. and A. ffrench (2011[29]), "Context validity", in Taylor, L. (ed.), *Examining Speaking: Research and Practice in Assessing Second Language Speaking*, UCLES/Cambridge University Press, Cambridge, p. 164.

The stage of **phonological encoding** involves the retrieval of appropriate phonology from memory, and a manifestation of this cognitive process is **spoken fluency**. Common indicators of fluency are the amount of planning time required before producing an utterance in real time, the placement and length of hesitations, the length of utterances, and grammatical and collocational accuracy (use of chunks of speech) (Field, 2011[27]). Looking at hesitation and planning pauses as an example, the CEFR scale of "Spoken fluency" (Council of Europe, 2020, p. 142[11]) demonstrates a clear gradation between the levels (Table 3.7).

Table 3.7. CEFR spoken fluency scale

CEFR level	Fluency
C2	Can express themselves at length with a natural, effortless, unhesitating flow. Pauses only to reflect on precisely the right means to express their thoughts or to find an appropriate example or explanation.
C1	Can express themselves fluently and spontaneously, almost effortlessly. Only a conceptually difficult subject can hinder a natural, smooth flow of language.
B2	Can communicate spontaneously, often showing remarkable fluency and ease of expression in even longer complex stretches of language.
	Can produce stretches of language with a fairly even tempo; although they can be hesitant as they search for patterns and expressions, there are few noticeably long pauses. Can interact with a degree of fluency and spontaneity that makes regular interaction with users of the target language quite possible without imposing strain on either party.
B1	Can express themselves with relative ease. Despite some problems with formulation resulting in pauses and "cul-de-sacs", they are able to keep going effectively without help.
	Can keep going comprehensibly, even though pausing for grammatical and lexical planning and repair is very evident, especially in longer stretches of free production.
A2	Can make themselves understood in short contributions, even though pauses, false starts and reformulation are very evident.
	Can construct phrases on familiar topics with sufficient ease to handle short exchanges, despite very noticeable hesitation and false starts.
A1	Can manage very short, isolated, mainly pre-packaged utterances, with much pausing to search for expressions, to articulate less familiar words/signs, and to repair communication.
Pre-A1	Can manage very short, isolated, rehearsed utterances using gesture and signalled requests for help when necessary.

Source: Council of Europe (2020[11]), *Common European Framework of Reference for Languages: Learning, teaching, assessment - Companion volume*, © Council of Europe, Strasbourg, p. 142.

At Level Pre-A1, quasi-fluent output only occurs with short rehearsed utterances that are likely formulaic in nature ("Can manage very short, isolated, rehearsed utterances using gestures and signalled requests for help when necessary", p. 142). An A1 learner's spoken production includes a great deal of hesitation or pausing as they search their lexical and grammatical resources before articulating their utterance. At Level A2, although pausing and false starts are still evident, the learners are more capable of making themselves understood. They show minimal signs of fluency when producing short utterances on familiar topics. A shift occurs at Level B1, where learners begin to show signs of fluency, even though hesitation or pausing may still occur, particularly in extended output. It is at this level that learners are less reliant on the help of the interlocutor when they face a breakdown in communication but can overcome problems on their own because of their wider lexical and grammatical resources. Spoken fluency is evident at Level B2, particularly in extended and spontaneous speech. Level B2 learners may still need to search for lexical or grammatical features, but the resulting pauses or dysfluency markers tend to be infrequent and are not particularly distracting for the listener. The use of conventionalised chunks or fillers to gain planning time (e.g. you know, like, etc.) may also begin to appear around this level, which Hasselgreen (2005, p. 118[30]) points out "contributes to the impression of fluency". At Level C1, only utterances on "conceptually difficult subject[s]" (Hasselgreen, 2005, p. 144[30]) will interfere with naturally fluent speech, and speakers may use hesitation to achieve a particular rhetorical effect (De Jong, 2018[31]), while at Level C2, speech is natural and effortless.

Two aspects of the **phonetic encoding and articulation** processes can impose cognitive load on the foreign language speaker: "inadequate phonological representations in the mind and inability to adjust to unfamiliar articulatory settings" (Field, 2011, p. 108[27]). For this reason, it has been suggested that the emphasis should not be on the accuracy of pronunciation in terms of native-speaker norms but on **intelligibility** (Field, 2011[27]). The relationship between intelligibility and proficiency focuses on the amount of effort required of the audience or interlocutor to understand the message because of the speaker's phonological features (Council of Europe, 2020, p. 133[11]), the listener's characteristics, or both (Isaacs, 2008[32]). Intelligibility can be affected by a speaker's control over articulation of phonological and/or prosodic features (i.e. stress, rhythm and intonation) (Field, 2011, p. 95[27]; Council of Europe, 2020, p. 133[11]). At the lower levels, learners may be able to produce a small set of learned words and phrases intelligibly, but their lack of familiarity with the phonological and prosodic features of the language may lead to speech that is heavily influenced by their first language and may be difficult to understand without support from the interlocutor.

As learners become more proficient, their intelligibility increases because they have had more exposure to the phonological and prosodic features of the target language and they have had more opportunities to adjust their speech in response to feedback when faced with breakdowns in communication related to intelligibility. By Level B1, learners are considered "generally intelligible" although stress, rhythm and/or intonation may still be influenced by their first language, and they will mispronounce words or sound combinations with which they are less familiar (Council of Europe, 2020, p. 134[11]). However, when issues around intelligibility arise, learners at Level B1 and above have the linguistic capability to find other means of resolving possible miscommunications, such as reformulating an utterance using different words.

Learners at Level B2 are considered intelligible, but they may still make systematic errors in pronunciation. They can generally use their phonological knowledge to pronounce unfamiliar words reasonably accurately, and even if their first language accent is noticeable, it does not interfere with intelligibility. By Level C1, learners have a high degree of control over the phonological and prosodic features of the language, leading to highly intelligible speech. The occasional errors made by these learners are usually self-corrected. Learners at this level can also use varied prosodic features to convey meaning more precisely. Level C2 learners are highly intelligible, even if their accent is still influenced by their first

language. This level is also characterised by the ability to use prosodic features to express "finer shades of meaning" (Council of Europe, 2020, p. 134[11]). Although phonetic intelligibility typically increases with proficiency (as described above), when pronunciation is effectively taught, phonetic intelligibility can be established early on at lower levels of proficiency.

Self-monitoring occurs to ensure that the intended message is being communicated and/or any breakdown in communication is resolved through some form of repair strategy.[10] Monitoring is possible at any level of proficiency, whereas self-repair depends on and increases with proficiency (e.g. Van Hest (1997[33])).The cognitive demands on planning and articulation at the lower levels may limit the availability of cognitive resources for self-monitoring and repair so that these learners rely on the interlocutor or audience to identify errors or communication breakdowns and help with repair. The ability to monitor and repair first appears in the CEFR at Level B1, focusing more on the ability to identify an issue than to necessarily independently resolve the issue (Council of Europe, 2020, p. 70[11]). By Level B2, learners are able to correct errors they notice, although they are not necessarily catching all errors. By the C levels, both the identification and correction of errors is effective and increasingly effortless. Finally, the higher levels are also characterised by the ability to monitor for pragmatic effectiveness alongside linguistic accuracy (Field, 2011[27]).

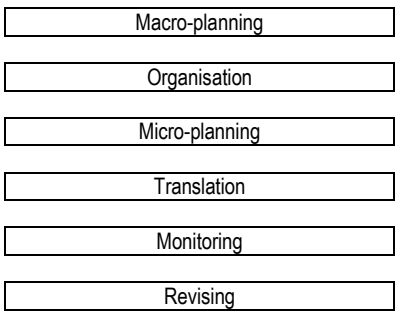
Annex 3.A provides an overview of the cognitive processes which would need to be employed to achieve the key aspects of the CEFR descriptors for speaking at each level.

Writing[11]

How do we write? A cognitive processing account of writing

Field's (2004[7]) model of L1 writing is based on information-processing principles, and as such will be used as the basis for a discussion of the cognitive processes employed in writing in the PISA Foreign Language Assessment. Figure 3.5 provides an overview of the stages within the model.

Figure 3.5. Stages of writing



Source: Field, J. (2004[7]), *Psycholinguistics: The Key Concepts*, Routledge, London.

The first three stages in the model relate to planning or conceptualisation. The first stage, macro-planning, occurs when the writer assesses the goals of the writing task in terms of the genre, the targeted reader, and level of formality needed. They then use this information along with their content knowledge to gather or generate ideas. Writers at this stage may also identify any constraints to the successful completion of the writing task, such as time limits or lack of content knowledge. In the next stage, organisation, the writer orders the ideas, determines the relationship between them and decides on their relative importance for achieving the goal identified in the previous stage. Next is micro-planning, which is when the writer focuses on the specific part of the text that is about to be written. Shaw and Weir (2007[34]) note that this planning

happens at two levels: the goal of the sentence within the paragraph and the goal of the paragraph in terms of how it fits within the larger piece of writing. Field (2005[35]) points out that the three stages of planning are abstract and within the mind of the writer. It is only at the next stage, translation, that the plans are converted into linguistic form. At this stage, the writer is making lexical and grammatical choices in order to achieve the macro-planning goals and ensure the text is functionally appropriate (Shaw and Weir, 2007[34]). The skilled writer will also include elements to support the development of a coherent and cohesive text (Shaw and Weir, 2007[34]). Monitoring for accuracy occurs both at the word or sentence level in terms of ensuring accurate linguistic forms and conventions are followed, but also at the discourse level in terms of ensuring the text is developing in line with the main goals determined at the macro-planning stage, and that each part of the text supports the argument being developed. As a consequence of monitoring, the writer will engage in revising and make changes to anything that is deemed to be incorrect or not achieving the writer's intentions. At this stage, writers may circle back to the planning stages if issues with ideas or their relative importance are identified (Field, 2004[7]).

How do foreign language learners write? Cognitive processes across CEFR proficiency levels

The cognitive demands that learners can cope with when writing will vary by proficiency level. The CEFR provides some indication of these expectations. It conceptualises writing as either written production (see Table 3.8) or written interaction (see Table 3.9).[12] Written production covers transactional and evaluative types of writing (e.g. reports, essays and creative writing). In contrast, written interaction focuses on interactive writing either for interpersonal reasons or the provision of information (e.g. notes, messages and correspondence, respectively).

As can be seen in Table 3.8 and Table 3.9, learners at Level Pre-A1 primarily write words usually about themselves, which may be required when filling in a form. This is only likely to require learners at this level to engage in translation. At Level A1, learners write simple sentences, phrases and clauses that are again personally relevant and primarily rely on translation, with very limited planning, monitoring and revising taking place. It is not until Level A2 that learners begin to write extended texts, such as simple paragraphs or letters. It is at this level that macro-planning, organisation and micro-planning are more explicitly involved (i.e. there are no planning descriptors below Level A2 in the CEFR). At Level B1, there is reference to straightforward organisational patterns in the written production scale (i.e. linear sequence) and getting across one's point in the written interaction scale. This suggests that learners are better able to engage in the planning stages, and, in particular, are becoming more aware of the reader and the need to consider this person when writing. Then at Level B2, learners are able to synthesise information from multiple sources as well as integrate or show relationships between their points and the points of others, which suggests a more sophisticated ability to micro- and macro-plan. By Level C1, learners are able to adjust their writing to express ideas precisely, making subtle distinctions in style, and demonstrating an understanding of the writer-reader relationship. By Level C2, learners are able to write highly effective texts that are coherent and cohesive.

Table 3.8. CEFR overall written production scale

CEFR level	Overall written production
C2	Can produce clear, smoothly flowing, complex texts in an appropriate and effective style and a logical structure which helps the reader identify significant points.
C1	Can produce clear, well-structured texts of complex subjects, underlining the relevant salient issues, expanding and supporting points of view at some length with subsidiary points, reasons and relevant examples, and rounding off with an appropriate conclusion. Can employ the structure

CEFR level	Overall written production
	and conventions of a variety of genres, varying the tone, style and register according to addressee, text type and theme.
B2	Can produce clear, detailed texts on a variety of subjects related to their field of interest, synthesising and evaluating information and arguments from a number of sources.
B1	Can produce straightforward connected texts on a range of familiar subjects within their field of interest, by linking a series of shorter discrete elements into a linear sequence.
A2	Can produce a series of simple phrases and sentences linked with simple connectors like “and”, “but” and “because”.
A1	Can give information about matters of personal relevance (e.g. likes and dislikes, family, pets) using simple words/signs and basic expressions. Can produce simple isolated phrases and sentences.
Pre-A1	Can give basic personal information (e.g. name, address, nationality), perhaps with the use of a dictionary.

Source: Council of Europe (2020[11]), *Common European Framework of Reference for Languages: Learning, teaching, assessment - Companion volume*, © Council of Europe, Strasbourg, p. 66.

Table 3.9. CEFR overall written interaction scale

CEFR level	Overall written interaction
C2	Can express themselves in an appropriate tone and style in virtually any type of formal and informal interaction.
C1	Can express themselves with clarity and precision, relating to the addressee flexibly and effectively.
B2	Can express news and views effectively in writing, and relate to those of others.
B1	Can convey information and ideas on abstract as well as concrete topics, check information, and ask about or explain problems with reasonable precision.
	Can compose personal letters and notes asking for or conveying simple information of immediate relevance, getting across the point they feel to be important.
A2	Can compose short, simple formulaic notes relating to matters in areas of immediate need.
A1	Can ask for or pass on personal details.
Pre-A1	Can convey basic information (e.g. name, address, family) in short phrases on a form or in a note, with the use of a dictionary.

Source: Council of Europe (2020[11]), *Common European Framework of Reference for Languages: Learning, teaching, assessment - Companion volume*, © Council of Europe, Strasbourg, p. 82.

During the translation stage, the grammatical and lexical requirements at different CEFR levels will often be linked to the functional requirements of tasks. As learners become more proficient, they are better able to flexibly meet the challenges of the task expectations because they have the necessary linguistic resources. Table 3.10 provides an overview of writing functions by CEFR level.

Table 3.10. Overview of writing functions by CEFR proficiency level

CEFR level	Overview	Selected examples
C2	Learners at this level have a deeper appreciation of the full range of functions. They can demonstrate sensitivity to their audience and are stylistically proficient.	Reviewing Justifying Analysing Judging priorities
C1	Learners at this level are able to use an appropriate and consistent register when meeting the functional requirements.	Making and justifying recommendations Evaluating a situation

CEFR level	Overview	Selected examples
	They should have an increasing awareness of how to express functions in stylistically different ways.	Hypothesising Advising
B2	Learners at this level are expected to deal with functions with a greater degree of range and accuracy.	Arguing a point of view Offering an explanation Expressing attitudes Comparing
B1	A slightly more demanding understanding of functions is required in terms of length, genre and language skills.	Describing future or imaginary situations Describing simple processes Drawing simple conclusions and making recommendations
A2	Basic functions relating to personal information, everyday activities and simple social interaction can be carried out.	Asking for and giving personal details Making and responding to offers and suggestions Expressing preferences, likes and dislikes

Source: Shaw, S.D. and C.J. Weir (2007[34]), *Examining Writing: Research and Practice in Assessing Second Language Writing*, UCLES/Cambridge University Press, Cambridge, pp. 125-127.

At the monitoring and revising stages, lower-level learners tend to lack the linguistic knowledge to identify their errors, or if identified, to correct them successfully. Because learners at Level Pre-A1 to A2 are using their cognitive resources at the translation stage, they spend little time monitoring or revising their writing (Field, 2004[7]). The ability to monitor and revise improves as learners reach Level B1, but learners are still narrowly focused on grammar and vocabulary rather than organising their ideas, and their ability to successfully revise is limited (Shaw and Weir, 2007[34]). It is not until learners reach Level B2 that they have the linguistic, pragmatic and sociolinguistic knowledge to engage in monitoring and revising in a more comprehensive manner, although with varying degrees of effectiveness. At Level C1, learners are expected to be able to effectively monitor and revise their texts to ensure they are meeting their intended goals. At Level C2, monitoring and revising is more efficient in that learners are able to quickly identify solutions to the issues they notice.

Annex 3.A provides an overview of the cognitive processes which would need to be employed to achieve the key aspects of the CEFR descriptors for writing at each level.

Of course, the recent advancements of technology in the area of writing aid tools, like machine translation, naturally raise the question of their impact in the context of foreign language writing instruction and learning. Box 3.1 looks at this issue.

Box 3.1. The impact of machine translation tools in foreign language writing

With improvements in machine translation (MT) in recent years, there is an increased interest in using translation tools in the classroom to support the development of learners' foreign language writing. At present, there appear to be two methods of use: learners use MT to convert a piece written in their first language into the foreign language and then edit the output, or write directly in the foreign language and use MT to support the editing process. MT is not widely used in educational contexts, however, because of concerns over its reliability in accurately translating a text into a foreign language. Lee (2020[36]) found that although MT can help learners develop lexical and grammatical range and accuracy, it is less useful for discourse level aspects of writing. MT use in the classroom can be motivational for students and improve self-directed learning. However, it can also lead to plagiarism and exposure to inaccurate lexis and grammar if not properly monitored (Garcia and Pena, 2011[37]). There is little evidence that MT is being used to replace foreign language writing instruction due to the limitations of the translations, but as technology gets better in time, it remains to be seen how MT may change the way individuals write or learn to write in a foreign language.

Notes

1. In the FLA Framework, “task” is used to refer to any purposeful language activity that language users and learners engage with; this includes, in the context of the foreign language test, any assessment items or exercises (e.g. a set of questions about a text), which are also referred to as (assessment) tasks.

2. In the PISA FLA Framework, the task factors also include what the PISA Reading Literacy Framework describes as text factors, i.e. the characteristics of the environment (including, but not limited to, characteristics of the textual support available to read or the auditory input reading text(s) that one engages with and their main features).

3. General cognitive processes involved in reading are considered universal and transferable across languages (Alderson, 1984[13]; Cummins, 1979[14]; Cummins, 1991[15]). It is the level of competence over those processes that may vary.

4. The illustration of the cognitive processes in the model and their descriptions in this text present these processes as consecutive stages, but this is an oversimplification for descriptive purposes only; it is acknowledged that these processes are dynamic and may take place in parallel.

5. Inferencing and building a mental model are represented as distinct stages in Khalifa and Weir (2009[2]) for the purposes of providing a working description of the reading process, which can guide the analysis of foreign language assessment tasks. It is acknowledged, however, that these two stages are interlinked; for example, in Kintsch's (1998[39]) text comprehension theory, inference generation participates in the construction of a mental model of a text.

6. The reading CEFR scales and descriptors are organised around the purpose for reading or specific reading genres/functions (Council of Europe, 2020, p. 53[11]) rather than cognitive processes. This approach to describing reading, however, does provide insight into the cognitive processes learners would be expected to employ depending on their proficiency level.

   For example, the “Reading for orientation” scale focuses on skimming or scanning texts to complete a specific task; thus it is linked to the “goal setter” and “expeditious reading” purposes in the reading model. The “Reading for information and argument” scale, on the other hand, focuses on more detailed understanding of texts, which would require more careful reading and a deeper understanding of the content of a text. The depth of understanding, particularly at the higher levels, maps to the higher-level cognitive processes in the reading model as readers are expected to identify “finer points of detail, including attitudes, and implied as well as stated opinions” (Council of Europe, 2020[11]). Both scales describe the task demands in terms of the type of information that can be understood or located and the characteristics of the texts by level, which does not feature in the reading model but will affect the activation of different cognitive processes.

   “Reading correspondence” and “Reading instructions”, on the other hand, are scales that emphasise the specific features of these text types and the reading behaviours associated with them or challenges faced by readers when reading them. For example, in reading correspondence, descriptors provide information on the extent to which a reader can understand colloquial or idiomatic language, which would map to several cognitive processes, such as lexical access and inferencing. The same holds true for the CEFR scales pertaining to listening, speaking and writing, which are discussed in the equivalent sections.

7. The CEFR scales do not make explicit reference to listening cognitive processes, except in the case of the scale “Identifying cues and inferring”. This scale includes aspects of self-monitoring and making inferences, and describes strategies to aid in identifying the relevance of information heard (i.e. discourse construction).

8. Assimilation is the process of making some vowels or consonants more similar to other nearby sounds.

9. Functions describe what we are trying to do through language (e.g. agreeing/disagreeing, persuading, suggesting, etc.).

10. Several theories have been put forward to describe the underlying processes involved in monitoring and repair. See Kormos (1999[38]) for a review of these theories.

[11]. As mentioned in Chapter 2, the PISA 2025 Foreign Language Assessment will not assess writing or develop writing assessment tasks. Therefore, only a concise description of writing is provided here for the sake of the PISA FLA Framework completion. As writing habits are changing and are expected to continue to change in the foreseeable future, so will the learning, teaching and assessment of writing. Future iterations of the PISA FLA Framework will, thus, require a comprehensive update of the skill of writing.

[12]. Online interactive communication was added in the latest set of CEFR scales (Council of Europe, 2020[11]), which captures multi-modal communication (including writing) that can occur either simultaneously or in a consecutive sequence.

References

Alderson, J. (2000), *Assessing Reading*, Cambridge University Press, Cambridge. [6]

Alderson, J. (1984), "Reading in a foreign language: a reading problem or a language problem", in Alderson, J. and A. Urquhart (eds.), *Reading in a Foreign Language*, Longman, London. [13]

Cohen, A. and T. Upton (2006), *Strategies in Responding to the New TOEFL Reading Tasks*, Educational Testing Service, Princeton, NJ, http://www.ets.org/toefl. [5]

Council of Europe (2020), *Common European Framework of Reference for Languages : Learning, teaching, assessment - Companion volume*, Council of Europe, Strasbourg, https://www.coe.int/en/web/education/-/common-european-framework-of-reference-for-languages-learning-teaching-assessment-companion-volume. [11]

Council of Europe (2018), *Common European Framework of Reference for Languages. Learning, Teaching, Assessment. Companion Volume with New Descriptors*, Council of Europe, Strasbourg, http://www.coe.int/lang-cefr. [12]

Cummins, J. (1991), "Interdependence of first- and second-language proficiency in bilingual children", in Bialystock, E. (ed.), *Language processing in bilingual children*, Cambridge University Press, New York. [15]

Cummins, J. (1979), "Linguistic Interdependence and the Educational Development of Bilingual Children", *Review of Educational Research*, Vol. 49/2, pp. 222-251, http://dx.doi.org/10.3102/00346543049002222. [14]

Cummins, J. (1976), "The influence of bilingualism on cognitive growth: a synthesis of research findings and explanatory hypotheses", Bilingual Education Project, Toronto. [16]

Cutler, A. and C. Clifton (1999), "Comprehending spoken language: a blueprint of the listener", in Brown, C. and P. Hagoort (eds.), *The Neurocognition of Language*, Oxford University Press, Oxford. [21]

De Jong, N. (2018), "Fluency in Second Language Testing: Insights From Different Disciplines", *Language Assessment Quarterly*, Vol. 15/3, pp. 237-254, http://dx.doi.org/10.1080/15434303.2018.1477780. [31]

Field, J. (2013), "Cognitive validity", in Taylor, L. and A. Geranpayeh (eds.), *Examining Listening: Research and Practice in Assessing Second Language Listening*, UCLES/Cambridge University Press, Cambridge. [20]

Field, J. (2011), "Cognitive validity", in Taylor, L. (ed.), *Examining Speaking: Research and Practice in Assessing Second Language Speaking*, UCLES/Cambridge University Press, Cambridge. [27]

Field, J. (2009), *Listening in the Language Classroom*, Cambridge University Press, Cambridge. [22]

Field, J. (2005), *Second language writing: a language problem or a writing problem?*. [35]

Field, J. (2004), *Psycholinguistics: The Key Concepts*, Routledge, London. [7]

Galaczi, E. and A. ffrench (2011), "Context validity", in Taylor, L. (ed.), *Examining Speaking: Research and Practice in Assessing Second Language Speaking*, UCLES/Cambridge University Press, Cambridge. [29]

Garcia, I. and M. Pena (2011), "Machine translation-assisted language learning: Writing for beginners", *Computer Assisted Language Learning*, Vol. 24/5, pp. 471-487, http://dx.doi.org/10.1080/09588221.2011.582687. [37]

Gernsbacher, M. (1990), *Language Comprehension as Structure Building*, Erlbaum, Hillsdale. [24]

Hasselgreen, A. (2005), *Testing the Spoken English of Young Norwegians: A Study of Test Validity and the Role of 'Smallwords' in Contributing to Pupils' Fluency*, UCLES/Cambridge University Press, New York. [30]

Hyland, K. (2002), "Teaching and Researching Writing"*, Applied Linguistics in Action Series*. [8]

Isaacs, T. (2008), "Towards defining a valid assessment criterion of pronunciation proficiency in non-native English speaking graduate students", *Canadian Modern Language Review*, Vol. 64/4, pp. 555-580. [32]

Khalifa, H. and C. Weir (2009), *Examining Reading: Research and Practice in Assessing Second Language Learning*, UCLES/Cambridge University Press, Cambridge. [2]

Kintsch, W. (1998), *Comprehension: A paradigm for cognition*, Cambridge University Press, Cambridge. [39]

Koda, K. (2005), *Insights into Second Language Reading: A Cross-linguistic Approach*, Cambridge University Press, New York. [18]

Koda, K. (1989), *The Effects of Transferred Vocabulary Knowledge on the Development of L2 Reading Proficiency*. [19]

Kormos, J. (1999), "Monitoring and Self-Repair in L2", *Language Learning*, Vol. 49/2, pp. 303-342, http://dx.doi.org/10.1111/0023-8333.00090. [38]

Lee, S. (2020), "The impact of using machine translation on EFL students' writing", *Computer Assisted Language Learning*, Vol. 33/3, pp. 157-175, http://dx.doi.org/10.1080/09588221.2018.1553186. [36]

Levelt, W. (1989), *Speaking: from intention to articulation*, MIT Press, Cambridge, MA. [28]

Oakhill, J. and M. Garnham (1988), *Becoming a Skilled Reader*, Blackwell, London. [25]

OECD (2019), *PISA 2018 Assessment and Analytical Framework*, PISA, OECD Publishing, Paris, https://dx.doi.org/10.1787/b25efab8-en. [3]

Perfetti, C. and L. Harris (2013), “Universal Reading Processes Are Modulated by Language and Writing System”, *Language Learning and Development*, Vol. 9/4, pp. 296-316, http://dx.doi.org/10.1080/15475441.2013.813828. [17]

Perfetti, C. and J. Stafura (2014), “Word Knowledge in a Theory of Reading Comprehension”, *Scientific Studies of Reading*, Vol. 18/1, pp. 22-37, http://dx.doi.org/10.1080/10888438.2013.827687. [1]

Schank, R. and R. Abelson (1977), *Scripts, Plans, Goals and Understanding*, Erlbaum, Hillsdale. [23]

Shaw, S. and C. Weir (2007), *Examining Writing: Research and Practice in Assessing Second Language Writing*, UCLES/Cambridge University Press, Cambridge. [34]

Sticht, T. and J. James (1984), “Listening and reading”, in Pearson, P. et al. (eds.), *Handbook of Reading Research*, Longman, New York. [10]

Urquhart, A. and C. Weir (1998), *Reading in a Second Language: Process, Product and Practice*, Longman, New York. [4]

Van Dijk, T. and W. Kintsch (1983), *Strategies of Discourse Comprehension*, Academic Press, New York. [26]

Van Hest, E. (1997), “Self-repair as a measure of L2 proficiency”, University of Jyväskylä. [33]

Weir, C. and H. Khalifa (2008), “A cognitive processing approach towards defining reading comprehension”*, Research Notes*, No. 31. [9]

Annex 3.A. Overview of the cognitive processes employed at each CEFR level

Annex Table 3A.1. The relationship between CEFR levels and cognitive processes employed for reading

CEFR level	Learners are likely employing
C2	All cognitive processes, including intertextual representations.
C1	All cognitive processes with minor limitations related to text content that may require rereading or reference tools to fully comprehend, with a good level of understanding, and in particular recognising implicit or implied meaning.
B2	Higher-level reading processes up to and including text-level representation with some challenges when reading outside their field of specialisation. First sign that learners can independently and successfully employ the goal setter for different reading purposes.
B1	Higher-level processes such as building a mental model, signs that learners can begin to build a text-level representation provided the text is clearly signposted (i.e. straightforward).
A2	Lower-level processes including some very basic world-level inferencing.
A1	Lower-level processes up to establishing propositional meaning. Understanding relies on familiar content.
Pre-A1	Word recognition and lexical access.

Source: Authors.

Annex Table 3A.2. The relationship between CEFR levels and cognitive processes employed for listening

CEFR level	Learners are likely employing
C2	All cognitive processes.
C1	All cognitive processes with minor limitations related to content or accent, but learners have the capacity to self-monitor and engage in appropriate strategies to overcome comprehension difficulties.
B2	Higher-level listening processes employed with some limitations when listening to poorly structured discourse or unfamiliar accents.
B1	First indication learners are engaging in higher-level processes, such as enriching meaning to understand the intent of spoken discourse; however, this relies on utterances being well-structured.
A2	Primarily lower-level processes.

A1	Lower-level processes of decoding, lexical search and parsing. Understanding relies on modified input.
Pre-A1	Decoding and lexical search, relying heavily on modified input in terms of reduced speech rate, increased pausing, use of high frequency, familiar vocabulary.

Source: Authors.

Annex Table 3A.3. The relationship between CEFR levels and cognitive processes employed for speaking

CEFR level	Learners are likely employing
C2	All cognitive processes with a high degree of control and flexibility.
C1	All cognitive processes, including the ability to monitor and correct pragmatic errors flexibly.
B2	All cognitive processes with increased control over them, which allows for more precise utterances.
B1	All cognitive processes, but with limitations. First indication learners are able to sustain extended discourse without significant rehearsal or preparation (fluency) and more cognitive resources available for self-correction.
A2	Grammatical, phonological and phonetic encoding. They have limited resources available for self-correction.
A1	Grammatical, phonological and phonetic encoding with limitations in terms of familiarity of context and time to prepare an utterance.
Pre-A1	Phonological and phonetic encoding.

Source: Authors.

Annex Table 3A.4. The relationship between CEFR levels and cognitive processes employed for writing

CEFR level	Learners are likely employing
C2	All stages with a high degree of control and flexibility.
C1	All stages of the writing model, including the ability to monitor and correct pragmatic errors flexibly.
B2	More balanced control over all stages of the writing process in order to achieve the intended impact on the reader.
B1	Increased emphasis on planning stages and monitoring and revising stages.
A2	Micro- and macro-planning is more evident because of the production of extended texts.
A1	Translation, with limited planning, monitoring and revising.
Pre-A1	Translation and orthographic control.

Source: Authors.

4 Assessing foreign language skills

This chapter describes how learner factors are accounted for in the PISA Foreign Language Assessment and how the test will elicit the cognitive processes described in Chapter 3 through tasks that reflect appropriate linguistic and contextual parameters. In doing so, this chapter outlines how reception activities (reading and listening comprehension) and spoken production will be assessed in the PISA 2025 cycle. Moreover, the chapter provides an overview of the different considerations the project is taking into account to ensure that the test enables fair and objective comparisons of language proficiency across school systems.

Chapter 3 outlined the conceptual framework for foreign language proficiency. The concepts in the framework must, in turn, be represented in tasks[1] and questions in order to measure students' proficiency in reading, listening to, and speaking, a foreign language.[2]

Foreign language proficiency is the outcome of two sources of influence: learner characteristics and task characteristics (including, but not limited to, characteristics of the text available to read or the auditory input, and aspects related to the goals or reasons for a particular activity). This chapter describes how learner factors other than proficiency in a foreign language are accounted for in the Programme for International Student Assessment (PISA) Foreign Language Assessment (FLA). It also describes how the test elicits the cognitive processes described in Chapter 3 through tasks that reflect appropriate linguistic and contextual parameters, as illustrated in the Common European Framework of Reference for Languages (CEFR). Other elements of test validity will be discussed in Chapter 6, which outlines how the PISA FLA is scored and how it is linked to the CEFR levels.

In accordance with PISA's aims, the two most important considerations in designing the PISA 2025 Foreign Language Assessment are: 1) to ensure broad coverage of the typical tasks and environments that are representative of foreign language use in school and wider society; and 2) to represent a range of task difficulty, aligned to the CEFR, by manipulating task demands as well as linguistic and contextual characteristics of the environment.

In the first iteration of the PISA Foreign Language Assessment, which will be part of the PISA 2025 cycle, English will be the only foreign language assessed. The FLA is an optional component of PISA, and having a minimum of participating countries is necessary to enable meaningful international comparisons as well as to make an assessment affordable. There were not enough countries interested in assessing another foreign language in the first cycle, though more languages can be included in future.

The first cycle will also limit the assessment to reception (listening and reading) and to spoken production. Therefore, the remainder of this chapter will focus on reading, listening and speaking only. Although the ability to communicate in writing is an important skill, particularly in light of the increasing use of digital technologies to interact and share one's point of view with others globally, often through some form of text, writing will not be assessed in the 2025 iteration. However, this decision does not preclude it from being included in future.

Addressing learner factors that influence proficiency

PISA is an international assessment of 15-year-old students, who differ widely in their background knowledge, in their native languages and languages of instruction, as well as in how proficient they are in English, the foreign language being assessed.

In developing tasks for the foreign language component, care is therefore taken to ensure that the relative difficulty of tasks is not dependent on group characteristics such as cultural background or gender. Content created for specific tasks is selected and developed from source material appropriate for the cohort. This appropriacy is determined both in relation to unifying cohort factors, such as age and enrolment in education, and with an awareness of the wide range of life experiences of candidates at an individual and national cohort level. Topic treatment is monitored throughout the task development process to avoid bias. This monitoring includes both expert reviews of all tasks involving representatives from all participating countries and statistical analyses.

In addition, a background questionnaire allows for the investigation of some of the factors, such as students' attitude toward the foreign language, self-efficacy and parental knowledge of the foreign language, which are expected to relate to proficiency (see Chapter 5). Another learner factor that is highly relevant in the context of an international assessment is the language distance between a student's first language (L1) and the foreign language being assessed (see Box 4.1). The questionnaires

will collect information on language(s) of instruction and on students' personal language background, and therefore enable to use a pre-existing classification of language distance to calculate students' language distance from English and to explore the relationship with their proficiency in English. However, these analyses would present all the limitations described in the box below.

Box 4.1. The effect of first language and foreign language distance on foreign language learning

Language transfer theory suggests that learning a foreign language may be easier and potentially more efficient if a learner's first language (L1) shares features with the foreign language being learned (Gass S.M. and Selinker, 1983[1]). Corder (1981, p. 101[2]) contends that when the L1 and foreign language are alike, learners "pass more rapidly along the developmental continuum" because they identify the similar L1 features of the foreign language more readily. An implication of this is that languages that do not share many similar features will be more difficult to learn and potentially take longer to learn. This notion has received popular support from the Defence Language Institute's School of Language Studies,[1] who categorise languages by learning difficulty based on their substantial experience teaching native English speakers these foreign languages. The languages are categorised based on linguistic and/or cultural differences from English. Although this categorisation is based on English speakers learning other languages, Chiswick and Miller (2005[3]) investigated whether it holds in the other direction in terms of the English language proficiency of immigrants to the United States and Canada, and have found parallels (Chiswick and Miller, 2005[3]).

Very few studies, however, have shown a clear relationship between language distance and learning success. This is primarily because it can be difficult to isolate language distance from other variables that contribute to language learning, such as language aptitude, exposure, motivation, parental influences, etc. Additionally, there is no one simple way to define language distance. For example, linguistic features can be traced back and grouped by language families or languages can be grouped according to the presence of similar features (e.g. cognates) or how much a speaker from one language understands a speaker from another language. Elder and Davies (1998[4]) found that when using these different methods, languages were classified differently – either more or less distant. Moreover, different aspects within the same language may vary in how similar they are to the target language. A language could share similar phonological or syntactic systems but differ in orthographic, which is to say, written representation.

When investigating language distance and learning, the relationship is not always straightforward. In fact, it has been found that some features are easier to learn when the first language does not share them (Odlin, 1989[5]). If shared features operate differently in the foreign language than in the L1, learners must engage in some form of reinterpretation of the feature in the foreign language, which can cause a learning difficulty (Lardiere, 2009[6]). This negative transfer is often the focus of research, particularly in the development of morphosyntactic and phonological features in the foreign language (Foley and Flynn, 2013[7]).

In terms of reading and listening, the differences between the L1 and foreign language writing systems have been found to influence reading processes (Koda, 2005[8]). These differences tend to have an impact on word recognition, making this a slower process and reducing reading fluency; but they do not appear to have a significant impact on overall comprehension (Koda, 2005[8]). Syntactic differences between languages can affect syntactic parsing, and differences in discourse features can affect how learners build mental representations of a written or aural text, but again, it is unclear whether these differences interfere with comprehension or simply processing speed (Koda, 2005[8]). Rost (2002[9]) indicates that for listening, the first language can affect lexical segmentation strategies in the foreign language. English relies on a "metrical segmentation strategy", which leads listeners "to assume every

strong syllable is the onset of a new word" (Rost, 2002, p. 96[9]). The extent to which the L1 shares this strategy will make input decoding easier. Lindgren and Muñoz (2013[10]) investigated a number of variables that influence listening and reading comprehension, with language distance as one variable. Their findings indicate that language distance, based on cognates (lexical similarity between languages), was the strongest predictor of reading and listening scores of the different variables investigated, but that the amount of variance it predicted was still relatively small. Their results may also have been influenced by the characteristics of the reading and listening tasks and the way the researchers determined language distance (i.e. lexical similarity).

Elder and Davies (1998[4]) found that the tasks designed to investigate linguistic distance can determine how much influence the L1 has, in that tasks measuring accuracy versus communicative efficiency can produce different results. This reinforces the importance of using a variety of tasks that tap into different aspects of language knowledge when investigating language distance.

Finally, Ringbom (1987[11]) found that although the first language does have some influence on learning a foreign language, this influence is nuanced and may change throughout the learning development cycle, coming in and out of prominence at different stages. Typically L1 influence is strongest at the early stages of learning, at the lower proficiency levels and in more communicative-focused tasks (Ringbom, 1987[11]). Therefore, as learners become more proficient, the effect of the first language on learning the foreign language may be less significant.

1. See Hardison et al. (2012[12]) for a description of the English as a second language (ESL) programme.

Finally, it is important to ensure that the target knowledge or skill for each task is not confounded by the ability to understand instructions or the familiarity with computer-based language tests. Test navigation interface and response formats are designed to avoid competing cognitive demands during the assessment. Students are given the opportunity to practice all response formats in a short familiarisation section preceding the test, which is designed to help them navigate the test interface (common to all PISA tests) and with the response formats (some of which are specific to each test). This familiarisation section is translated into the local language of instruction. Its goal is to ensure that all students are familiar with the task types and other environmental features before they sit the test. All in-test instructions for students on how to answer a question are provided in English (the foreign language assessed in 2025). Care is taken to ensure test instructions are accessible to all students, irrespective of their language ability, by including visual cues (icons, pictures) and using the simplest vocabulary and grammar. For consistency, British spelling conventions are used for all instructions. In the speaking and listening sections, instructions are provided through multiple channels - on the screen and as an audio stream.

Assessing reading in a foreign language

In constructing the reading component for the foreign language assessment, care is taken to ensure sufficient coverage of the cognitive processes involved in reading and the CEFR descriptors corresponding to different ability levels in order to draw conclusions about reading competence. As illustrated in Chapter 3, the mastery of reading processes is closely related to a progression in readers' proficiency: language learners at lower levels of proficiency will employ a more narrow set of processes when reading, focused on establishing literal meaning at the sentence level. Indeed, their limited lexical and grammatical knowledge may interfere with comprehension beyond the local level. More proficient language learners have mastered a wider range of processes because they have a greater breadth and depth of linguistic knowledge to engage in meaning construction beyond the literally stated one. They can draw inferences for assumed or missing information to identify attitudes and opinions, as well as trace the relationship between ideas in a text to identify an author's purpose and build a mental model of it. The most proficient

readers engage with a wide range of texts independently, and can efficiently adjust the way they read depending on the type of text and specific purpose for reading; they can recognise the hierarchical structure of a whole text and make connections between several texts.

As the CEFR's focus is on language use, reading scales provide information on task demands at each ability level rather than explicitly focusing on cognitive processes. That is, the CEFR descriptors for reading provide guidance on the purpose for reading, text characteristics, subject domains and depth of understanding expected. Learners at lower proficiency levels can understand concrete information in short, simple texts on familiar topics, and as their proficiency increases, the range of text types and topics that they can read effectively and efficiently increases.

Text characteristics and factors that affect reading processes

Text characteristics, such as text type, domain, length, etc., are controlled to ensure that the assessment tasks reflect appropriate linguistic and contextual parameters for the target level of proficiency, in accordance with the corresponding CEFR descriptors. A close analysis of the CEFR descriptors (see Annexes 4.A, 4.C and 4.E) and consideration of research, therefore, guides the development and selection of texts and items.

Reading comprehension involves a complex interaction between the features of the text, task and learner. The CEFR provides general information on text characteristics, which do require some interpretation. Therefore, a great deal of effort has been focused on resolving this ambiguity by identifying textual features that affect text difficulty for foreign language learners.[3] However, Alderson (2000[13]) points out that a drawback of this pursuit is that the interaction between the reader and text is often not considered. That is, reader variables such as their motivation to read, world knowledge or topic familiarity can affect comprehension (Alderson, 2000[13]), but are not taken into account in this approach. Therefore, for the PISA 2025 Foreign Language Assessment, care is taken to select texts not only based on textual features but also on the characteristics of the test takers (international, 15-year-olds, in education). This cohort of test takers, though sharing age and a current context of education, spans an extensive range of cultures and backgrounds. Topics and treatment of topics across the test will be carefully controlled to ensure that all content is accessible to students from diverse backgrounds and avoids any potential bias.

The remainder of this section will focus on the textual features controlled for during test development because they have been shown to affect text difficulty. It also considers factors that can improve domain coverage across the tasks developed for the reading portion of the PISA Foreign Language Assessment.

Reading texts are adapted from authentic source texts that are likely to be familiar, relevant and reflect the interests of 15-year olds. A range of texts written for different purposes are included: to inform, to convince, to convey feelings or emotions, to entertain or to keep in touch (phatic) (Jackobson, 1960[14]). As the CEFR descriptors in Table 3.1 (in Chapter 3) indicate, the type of text a learner can understand at lower proficiency levels would primarily be informational in nature, whereas at the higher levels, texts designed to persuade or convince would be appropriate.

Other ways of describing or classifying texts are by genre or discourse mode, which loosely is defined as describing how a text is organised to help readers navigate through it. Urquhart (1984[15]) points out that a text's difficulty can be affected by how it is organised with some rhetorical patterns leading to better recall. Texts have been classified in the PISA Foreign Language Assessment of reading as description, narration, exposition, argumentation, instruction and transaction, which is in line with the description of text types from the Reading Framework for PISA 2025 and Alderson et al. (2004[16]). Research indicates that narrative texts are easier to process in a foreign language than expository texts because of the universality of the form: narrative texts do not require any formal training to understand the text structure and tend to be propositionally less dense than expository texts (Koda, 2005[8]). Therefore, descriptive and narrative text types are most appropriate for the lower proficiency levels and exposition and argumentation for the

higher levels. Every effort is made to represent this range of diversity in the PISA Foreign Language Assessment of reading.

The CEFR descriptors also describe the context of language use, which is categorised into four domains: personal, public, occupational and educational, which represent "broad sectors of social life" (Council of Europe, 2001, p. 10[17]). These domains are included in the Reading Framework for PISA 2025 and are referred to as "situations". The personal domain relates to areas of personal interest (e.g. friends, family, home life, hobbies, etc.); the public domain comprises situations in wider society (e.g. public transport, restaurants, hospitals, etc.); the occupational domain focuses on work-related activities; and the educational domain on educational activities. Within each domain learners would encounter different types of texts; for example, in the personal domain, personal emails, brochures, junk mail and notes would be typical. Public announcements, notices and timetables for public transport would fall within the public domain. Business emails or letters, job descriptions and instructional manuals would be considered part of the occupational domain, and academic articles and textbooks would be part of the educational domain. The CEFR acknowledges that these categories are not necessarily mutually exclusive, and texts may fall under more than one domain depending on their specific characteristics or people involved (Council of Europe, 2001[17]).

The CEFR descriptors provide guidance on the domains appropriate at different proficiency levels. As PISA is testing 15-year-old students still in school, occupational texts will not be included. The texts at the lowest levels of proficiency included in the FLA relate to the personal or public domains, while texts set within the educational domain will be added at higher levels of proficiency. The topics are carefully selected to avoid subjects that 15-year-olds would not be expected to fully understand because of their age. This is particularly relevant for the higher levels as the texts become more complex.

As the CEFR is designed to describe what learners can do at different proficiency levels, it tends to do this in terms of functions within the domain. Functions describe what we *do*, or try to do, through language (e.g. agreeing/disagreeing, persuading, suggesting, etc.). The functional notional analysis done by Van Ek and Trim (1998[18]; 1998[19]; 2001[20]) and Wilkins (1973[21]; 1976[22]) contributed to the development of the CEFR and can be seen in the CEFR's action-oriented, "can-do" descriptors. The range of functions learners can cope with at the different proficiency levels is quite large, but it is often how they are presented linguistically that determines the level of difficulty. Green (2012[23]) analysed functional progression and found that learners increasingly expand the contexts of function use from a very limited repertoire of functions that are context-specific and restricted to expressing personal information to a wider range of functions that are used to maintain social interactions in a variety of contexts. A writer can select from a range of functional exponents to convey meaning to the reader at a micro level. These functions and their exponents, taken collectively, can contribute to the overall purpose of the text. For example, at lower levels more direct exponents such as *in my opinion, I think that,* could be used to give an opinion. At higher levels, less direct exponents might be used to give an opinion, such as *without doubt, it seems clear that.* This progression in functions is taken into account in source text selection and task development. An overview of the textual features indicative of proficiency level for reading is shown in Table 4.1.

Table 4.1. Overview of textual features indicative of level for reading texts

CEFR target level	Textual features
Low (Pre-A1 and A1)	**Overall purpose of text**: To inform **Discourse modes**: Descriptive, narrative **Domains**: Personal, public **Functions**: Learners at these levels are expected to comprehend texts by interpreting functional exponents expressed directly within the text, which are used to convey: personal information; information related to everyday activities and social interactions; likes/dislikes; and/or factual descriptions of activities.

CEFR target level	Textual features
Medium (A2, B1)	**Overall purpose of text**: To inform, to convey feelings and emotions **Discourse modes**: Descriptive, narrative, instructive, expository **Domains**: Personal, public, educational **Functions**: Learners at these levels are expected to comprehend texts by interpreting functional exponents expressed directly and less directly within the text, which are used to convey: personal information; information related to everyday activities and social interactions; opinions; requests for information and questions about preferences feelings and emotions and the responses to such requests and questions; factual descriptions of people and their lives; and/or rules and regulations.
High (B2 and above)	**Overall purpose of text**: To inform, to convey feelings and emotions, to entertain and delight, to convince or persuade **Discourse modes**: Descriptive, narrative, instructive, expository, argumentative **Domains**: Personal, public and educational **Functions**: Learners at these levels are expected to comprehend texts by interpreting functional exponents expressed directly and indirectly within the text in a range of registers and topics, which are used to convey: personal and more abstract information; attitudes and opinions, reasoned argument and criticisms of arguments; and/or detailed descriptions of experiences, events and places expressed in a literary or entertaining way.

Source: Cambridge Assessment English's original work, based on Khalifa, H. and C.J. Weir (2009[24]), *Examining Reading: Research and Practice in Assessing Second Language Learning*, UCLES/Cambridge University Press, Cambridge, pp. 116-117 and pp.105-110.

As learners become more proficient, they are able to cope with increasingly longer and more complex texts. Therefore, the length of texts does not exceed a single sentence at the lowest levels of proficiency (Pre-A1); at Levels A1-A2, they are limited to very short texts relating to everyday, familiar contexts of a concrete nature, with longer texts, composed of multi-paragraph passages introduced at Level B1. At the highest levels, texts become increasingly longer and more complex or abstract. Additionally, the propositional density increases from Level B2 upwards.

Lexico-grammatical characteristics of reading texts

The lexical and grammatical characteristics of texts are often considered key indicators of difficulty for reading texts. A number of studies have investigated whether one contributes more to text comprehension than the other (Choi and Zhang, 2018[25]). Although the results are mixed, it is clear that both grammar and vocabulary play an important role in comprehension, particularly at lower proficiency levels, when learners are less able to engage in higher-level cognitive processes to determine meaning because of gaps in their linguistic resources (Choi and Zhang, 2018[25]). The grammar and vocabulary competences in the CEFR are underspecified and intentionally general because the CEFR is designed to be language neutral. Grammatical and lexical competences for each ability level must be specified by language in reference level descriptors (RLD) because linguistic development will vary by language (Council of Europe, n.d.[26]). For English, a consortium of institutions has been working together to develop RLDs for grammar and vocabulary, which can provide some insight into which grammatical and lexical features learners can cope with at different levels of proficiency (Cambridge University Press, n.d.[27]).

Texts at the lowest levels are comprised of short, simple sentences, whereas as proficiency increases, texts include a wider range of verb forms, sentence patterns and grammatical complexity.

In terms of vocabulary, a number of research studies have suggested that learners must know between 95% and 99% of vocabulary in a text in order to engage in higher-level reading processes (Hu and Nation, 2000[28]; Nation, 2006[29]; Schmitt, Jiang and Grabe, 2011[30]). In light of this, careful consideration is given to the vocabulary within texts, in particular, the vocabulary that is directly associated with answering an item. In some cases, higher-level vocabulary may be present in a text to tap into students' ability to cope with unknown vocabulary; however, accurate understanding is not required to answer the item successfully. Texts used for lower proficiency levels contain high-frequency vocabulary related to everyday

topics in the personal and public domains. The breadth and depth of vocabulary present in texts expand as the targeted proficiency level increases.

Variety of English

Finally, as students in different educational contexts are likely to be learning different varieties of English (typically either British or American), when selecting texts for the PISA test, care has been taken to avoid disadvantaging students. For students at lower proficiency levels, this means using a more neutral, unmarked form of English. In practice, employing a neutral form entails, wherever possible, the avoidance of lexis, grammar and spelling with marked differences in British and American varieties. Examples of such marked language include: lexis – "jumper" (British English); grammar – "have got" (British English); spelling – "color" (American English). With regards to spelling, in cases where the use of words with a differing British and American spelling is unavoidable in texts for lower levels, the British spelling is used as the default, for consistency (e.g. "centre" rather than "center"). It is worth noting that there is no research to suggest that it would interfere with comprehension. In fact, there is evidence that both varieties of English co-exist in many educational contexts (Modiano, 1996[31]). At the higher levels, texts are typically taken from authentic sources and may, therefore, be more marked as British or American English by containing a range of features associated with either variety of English. At these higher levels, a balanced approach to text selection is taken: higher-level learners have typically had enough exposure to both varieties of English to be able to deal with them equally well (Modiano, 1996[31]).

Overview of test tasks, levels and processes

The reading test consists of a number of discrete assessment tasks, and longer tasks corresponding to multiple items sharing a common stimulus text or image (as is usual in PISA assessments). To aid with the description of reception activities and strategies, the CEFR provides an overall reading comprehension scale (see Chapter 3) alongside more specific scales for different forms of reading. Table 4.2 provides an overview of the CEFR reception scales targeted at each level in the PISA Foreign Language Assessment reading test through a range of task types.

Table 4.2. Overview of CEFR scales tested by proficiency level for reading

CEFR level	CEFR scales represented
Pre-A1	Overall reading comprehension Identifying cues and inferring
A1	Overall reading comprehension Identifying cues and inferring Reading correspondence
A2	Overall reading comprehension Identifying cues and inferring Reading correspondence
B1	Overall reading comprehension Identifying cues and inferring Reading correspondence Reading for information and argument
B2	Overall reading comprehension Identifying cues and inferring Reading for orientation Reading for information and argument
C1	Overall reading comprehension Identifying cues and inferring Reading for orientation Reading for information and argument

Source: Cambridge Assessment English's original work.

A detailed mapping of levels, tasks and specific descriptors is included in Annex 4.A. It maps the task type, CEFR level, task description, cognitive process, CEFR can-do statements and reading type. An extract from this mapping is presented in Table 4.3. Examples of the task types can be found in Annex 4.B.

Table 4.3. Extract from the mapping of reading tasks to CEFR level, cognitive process, CEFR can-do statements and reading type

Task type	CEFR level	Task description	Cognitive process	CEFR can-do statements	Reading type
Reading – Discrete definition with graphic	Pre-A1	**Input text**: Picture **Response format**: Three short sentences **Operation**: Select the sentence that defines the picture	Word recognition and lexical access	**Overall reading comprehension** Can recognise familiar words/signs accompanied by pictures, such as a fast-food restaurant menu illustrated with photos or a picture book using familiar vocabulary. **Identifying cues and inferring (spoken, signed and written)** Can deduce the meaning of a word/sign from an accompanying picture or icon.	Careful, local

Note: See Annex 4.A. for the complete mapping table.
Source: Cambridge Assessment English's original work, based on Council of Europe (2020[32]), *Common European Framework of Reference for Languages: Learning, teaching, assessment – Companion volume*, © Council of Europe, Strasbourg, pp. 54-60, and on Khalifa, H. and C.J. Weir (2009[24]), *Examining Reading: Research and Practice in Assessing Second Language Learning*, UCLES/Cambridge University Press, Cambridge, p. 70.

As can be seen in Table 4.3, the simplest task included in the assessment elicits information about decoding skills (word recognition and lexical access). Other tasks used at lower proficiency levels include matching words with pictures or definitions, multiple-choice, gapped sentence tasks with gaps corresponding to single words, and short texts with multiple-choice questions focused on specific detail. In order to assess students' ability to build a mental model and engage in meaning comprehension beyond literal understanding, texts with multiple-choice questions focusing on identifying main ideas, opinions, attitudes, the writer's purpose and making inferences are included. To assess their ability to create a text-level structure, gapped text tasks with entire sentences or passages removed and jumbled for students to match back to their original locations to create a coherent and cohesive text are used. At higher levels of proficiency, tasks that require students to compare and contrast the opinions expressed in different texts are used to assess students' mastery of higher-level reading processes such as creating an intertextual representation. For tasks at Level B1 and upwards, each stimulus is followed by multiple comprehension questions in order to ensure efficient use of testing time. Task management demands are kept at a minimum at the lowest levels of proficiency, whereas at the higher levels of proficiency tasks may include a time limit to ensure that students engage in expeditious reading.

Response formats: Reading

The response formats included in the assessment are selected to avoid method bias while maintaining construct coverage. Research indicates that different response methods can tap into different aspects of reading ability; therefore, including a range of response formats, if possible, is beneficial (Alderson, Clapham and Wall, 1995[33]; Graves et al., 1991[34]) (see also the PISA 2025 Reading Framework discussion on response formats).

Constructed response questions, however, are challenging to include in the PISA Foreign Language Assessment because foreign language skills do not necessarily develop in unison, and as a result, students' writing ability or familiarity with an English keyboard could significantly affect their ability to demonstrate reading comprehension. It is important to avoid this potential issue in the first cycle of the test; therefore, no constructed response formats are included.

Multiple-choice and matching formats were chosen because they are likely to be familiar to test takers. Multiple-choice questions are common in large-scale assessments because: they can be reliably marked, students are generally familiar with this format, they contribute to internal consistency, they can be effective in differentiating strong students from weak ones, and their level of difficulty can be adjusted, along with text features, to target a range of cognitive processes at different proficiency levels (Khalifa and Weir, 2009[24]). For example, multiple-choice questions can measure lower-level processes such as detailed understanding, requiring careful reading at the local level to establish propositional meaning. They can also measure higher-level processes such as building a mental model through careful reading at the global level. As such, this question format is used in the FLA at every CEFR level to elicit the cognitive process from word recognition to building a mental model.

Another format, which shares characteristics with the multiple-choice format, is matching where several items share the same options. In its simplest form, a student is presented with a column of pictures and a column of words and asked to match the words to the pictures (eliciting word recognition and lexical access). Typically, there are more words (options) than pictures (items) to reduce interdependence: if no distractors are present, a student who chooses one incorrect option will automatically get two incorrect items (Alderson, Clapham and Wall, 1995[33]). An advantage of the matching task is that it reduces the probability of guessing the correct answer by the number of options included (Brown and Hudson, 1998[35]).

As with multiple-choice questions, matching can elicit a number of cognitive processes but is ideal for several higher-level processes of interest. For example, a matching task at Level B2 to measure expeditious reading consists of four short texts on a related topic along with ten questions. Students are given five minutes to identify which text contains the answer to each question. Because similar information can be found in all four texts, students cannot simply match words (i.e. scanning) but must search for the answers, taking into account differences between texts. Readers must quickly decide which parts of a text to ignore or give only a small amount of attention to, and which to attend to and possibly read more carefully, which is a cognitively demanding task (Urquhart and Weir, 1998[36]).

Another type of matching task is the gapped text. Sentences (at Level B1) or paragraphs (at Level C1) are removed from a long text, jumbled and placed alongside distractors. Students must determine which sentence or paragraph goes into each gap. This task type requires an understanding of text coherence and cohesion as the reader must consider which information comes before and after the gap to determine which sentence/paragraph goes into the gap. The development of a mental model is necessary in order to follow the development of ideas at Level B1. In contrast, at Level C1, students must distinguish between major and minor supporting ideas to impose a hierarchical structure of the text. At Level C1, a matching task is included to elicit evidence of students' ability to build an intertextual representation. Such a task consists of four texts on a related topic with questions that place the texts in contrast; for example, which text shares the opinion of text X about Y? This requires students to compare the texts and recognise their similarities and differences. In summary, matching tasks are included in the FLA at the lower proficiency levels to elicit decoding, word recognition and lexical retrieval. At higher proficiency levels, they are used to assess expeditious reading, building a mental model, and creating a text-level structure and intertextual representation.

For multi-item, multiple-choice tasks that elicit careful reading, items are presented in the same order as in the text because this supports the type of reading the task is designed to test (Hughes, 2003[37]; Weir, 1993[38]). That is, ordering the items sequentially encourages the reader to read carefully in a linear manner, imposing a structural representation as they read (Khalifa and Weir, 2009[24]). For matching tasks,

items are presented in a jumbled order to encourage readers to first engage in expeditious forms of reading and then switch to careful reading at the local level to verify information (Khalifa and Weir, 2009[24]). All reading items are automatically scored.

Reading test forms

As the goal of PISA is to report at the population level, the entire item pool is assembled in multiple and partially overlapping forms, with each student taking only one form. To construct test forms, assessment units are assembled into longer sequences (or testlets), which combine different tasks at adjacent CEFR levels. Each testlet covers a broad range of cognitive processes, task types and reading types, within a relatively narrow range of task difficulty. A test form consisting of multiple testlets will therefore provide broad coverage of the reading processes and types of reading that are relevant at the particular proficiency levels included in the test. Table 4.4 provides approximate distributions of tasks by the cognitive process involved and testlet. At the stage of test development, all tasks are described in terms of the target process to ensure a balanced representation of these in the test.

Table 4.4. Distribution of reading tasks by cognitive process and proficiency level

CEFR level	Word recognition / lexical access	Syntactic parsing and Establishing propositional meaning at sentence level	Inferencing	Building a mental model	Creating a text-level structure	Creating an intertextual representation	Overall
Low (A1 and below)	40%	60%					25%
Medium (A2, B1)		35%	15%	50%			50%
High (B2 and above)				60%	20%	20%	25%

Note: Lower-level cognitive processes (to the left) are nested within higher-order processes (to the right).
Source: Cambridge Assessment English's original work, based on Khalifa, H. and C.J. Weir (2009[24]), Examining Reading: Research and Practice in Assessing Second Language Learning, UCLES/Cambridge University Press, Cambridge, p. 70.

The assignment of students to forms is adaptive, meaning that students' performance in the initial sections of the test determines, in part, the final assignment of students to forms. In particular, the adaptive design is expected to result in the majority of students being assessed, for most of the test, on items that broadly correspond to their level of proficiency. This will, in turn, ensure that students are assessed on reading processes and reading types that are represented in the CEFR proficiency descriptors, at their level of proficiency. To achieve this, the initial, non-adaptive section of the test is expected to correspond to approximately the median level of proficiency across participating countries, and students would move to testlets targeted at higher or lower levels depending on their responses to this initial section.

Assessing listening in a foreign language

Similar to the reading test, the listening test for the PISA Foreign Language Assessment is constructed to ensure good coverage of all processes identified in Chapter 3, and represent the CEFR descriptors for listening comprehension corresponding to different ability levels, so as to draw valid and reliable conclusions.

Task characteristics and factors that affect listening processes

A number of task characteristics – such as semantic and grammatical characteristics, speech rate and accent – are a function of the target proficiency level as described in the CEFR and research on listening comprehension. These features, therefore, are considered in the task design and selection process as described below.

Input text characteristics: The listening script

Listening scripts are typically sourced from authentic listening material, but modified in line with CEFR descriptors for each ability level. Although the use of unscripted listening tasks has merit on the grounds of authenticity (Wagner, 2013[39]), it cannot be assumed that learners sitting the PISA Foreign Language Assessment will have the same access or exposure to unscripted listening tasks in the learning environment, which could be a source of bias.

Although grammatical and lexical features of a script are often viewed as the main indicators of script difficulty, Field (2013[40]) points out that the semantic characteristics, or idea units, of the audio input contribute more to the cognitive demands placed upon the listener than linguistic features (Field, 2013, p. 121[40]). According to Chafe (1979[41]), idea units describe the pieces of information that contribute to the overall discourse representation constructed by the listener. Learners need to identify each idea unit by parsing the clause it appears in and tracing the links between idea units. As learners become more proficient, they can extract idea units in longer and more complex clauses and find it easier to link idea units within a larger discourse structure (Field, 2013[40]). They can also handle more abstract input provided it is contextualised appropriately (Schwanenflugel and Stowe, 1989[42]). Because the ability to process idea units effectively and comprehend abstract content only appears from Level B2 upwards, the semantic characteristics of spoken scripts chosen for FLA below this level are concrete and factual. At Level B1 there is a shift to slightly less concrete content, such as the identification of clearly stated attitudes or opinions, which are presented in well-organised and signposted scripts.

Other characteristics of audio input, such as listening purpose, domain, discourse modes and functions are controlled by proficiency level and support the semantic characteristics described above. Many of these features are similar to those described in the previous section on reading. At the lowest levels, the text purpose tends to be informational, focusing on the personal domain and on exchanging factual information about everyday topics. As students become more proficient, more persuasive and expressive purposes are employed, a wider range of domains and topics are included (i.e. educational, public), and a wider range of functions, such as expressing and identifying attitudes, and suasion are tested. Additionally, at higher proficiency levels, functions are realised in a greater variety of ways. A summary of the textual features indicative of each ability range for listening is shown in Table 4.5.

Table 4.5. Overview of textual features indicative of level for listening scripts

CEFR target level	Textual features
Low (A1 and below)	**Overall purpose**: To inform **Discourse modes**: Descriptive, narrative **Domains**: Personal, public **Functions**: Learners at these levels are expected to comprehend spoken texts by interpreting functional exponents expressed directly within the spoken text, which are used to convey: personal information; information related to everyday activities and social interactions; likes/dislikes; factual descriptions of activities; and/or requests for information and corresponding responses.
Medium (A2, B1)	**Overall purpose**: To inform, to convey feelings and emotions **Discourse modes**: Descriptive, narrative, instructive **Domains**: Personal, public, educational **Functions**: Learners at these levels are expected to comprehend spoken texts by interpreting functional exponents expressed directly and less directly within the spoken text, which are used

CEFR target level	Textual features
	to convey: personal information; information related to everyday activities and social interactions; opinions; requests for information and questions about preferences, feelings and emotions and the responses to such requests and questions; factual descriptions of people and their lives; rules and regulations; and/or decisions on a course of action.
High (B2 and above)	**Overall purpose**: To inform, to convey feelings and emotions, to entertain and delight, to convince or persuade **Discourse modes**: Descriptive, narrative, instructive, expository, argumentative **Domains**: Personal, public and educational **Functions**: Learners at these levels are expected to comprehend spoken texts by interpreting functional exponents expressed directly and indirectly within the spoken text in a range of registers and topics, which are used to convey: personal and more abstract information; attitudes and opinions, reasoned argument and criticisms of arguments; and/or detailed descriptions of experiences, events and places.

Source: Cambridge Assessment English's original work, based on Elliott, M. and Wilson, J. (2015[43]), *Context Validity in Examining Listening: Research and Practice in Assessing Second Language Listening,* UCLES/Cambridge University Press, Cambridge, pp. 203-210, and on Khalifa, H. and C.J. Weir (2009[24]), *Examining Reading: Research and Practice in Assessing Second Language Learning*, UCLES/Cambridge University Press, Cambridge, pp. 116-117 and pp. 105-110.

Lexico-grammatical characteristics of listening scripts

The grammatical and lexical characteristics of a spoken text can affect difficulty in terms of speed of lexical retrieval and parsing (Field, 2013[40]). The more familiar the syntactic structure is, the easier it will be to process. Consequently, short, simple utterances are featured at the lower CEFR levels, and an increasingly wider range of grammatical structures, tenses and aspects appear as students become more proficient. Similarly, the inclusion of high-frequency and familiar words that relate to personal and everyday topics will support lower-level students in their comprehension. As proficiency increases, students expand both the breadth and depth of lexical knowledge with comprehension of idiomatic language appearing from Level B2 and above.

Speech rate

One of the difficulties facing foreign language learners is that spoken input can be highly variable, and the spoken form of a word may not reflect its written counterpart, particularly in connected speech (Field, 2013[40]). In order to support lower-level learners, speakers will often reduce their speech rate to make it easier for learners to extract meaning. A slower speech rate is associated with increased pausing at word and sentence boundaries (Goldman-Eisler, 1968[44]) and, as such, aids test takers in lexical search by making word boundaries more distinct. This is particularly important in listening tests when the listener may not have control over the input and cannot rely on additional cues (e.g. facial expressions) to aid comprehension. Therefore, the speech rate has been modified at the lower proficiency levels to support students' developing decoding skills and lack of automatic processing capability at these CEFR levels (Field, 2013[40]). Great care, however, is taken during the recording of the audio to ensure that natural prosody and connected speech is maintained even when the rate of speech is slowed. Fewer adjustments to the speech rate are made when recording the audio for learners at higher proficiency levels.

Number of speakers

The number of speakers present in the audio input can also affect comprehension. Because of differences in speaking characteristics and the acoustic features of individual speakers, learners need to adjust to new voices (Field, 2013[40]; Brown and Yule, 1983[45]). This indicates that for lower-level learners, it is ideal if a maximum of two voices are used in any one recording and ideally, male and female voices are represented to make distinguishing between the speakers easier. An explicit effort is made to progressively vary the interactional patterns in the audio stimulus in line with proficiency in order to provide broad coverage of

authentic situations in which listening comprehension is required. For example, at lower proficiency levels, only informal dialogues are included. The question and answer format of these dialogues supports the identification of important information because the question highlights the content of the following turn. At higher proficiency levels, both dialogues and monologues are included. Additionally, at the lower levels, only two speakers are ever included, one male and one female, to ensure students can distinguish one speaker from the other and avoid confusion over who said what when answering questions. At Level B2 and upwards, up to three speakers may be included. Finally, to allow learners to adjust to the voices, questions are not asked about information contained in the first few seconds of a recording.

Variety of English

As discussed in the reading section, the written content for listening tasks (i.e. questions seen on screen) at lower proficiency levels will be presented in a more neutral variety of English, not marked overall as either British or American. At higher-level listening, the same approach will be adopted as with reading, with a balanced use of British and American English.

However, unlike reading, an important question concerns accents used in the delivery of audio input. There is a multiplicity of accents in English, with two main families of accents most likely encountered in international textbooks and learning material, these being standard British and standard North American. English language learners in different countries will have experienced different exposures to these two families of accents. Outside of these predominant families, students' familiarity with other English accents (including regional varieties) will be largely serendipitous, dependent on personal experiences. Also, in the context of English as a lingua franca, learners may have differing exposure to speakers of English from a range of first language backgrounds. Typically, students can cope with a range of different accents only at the highest levels of proficiency[4] (Field, 2013[40]).

Therefore, in an international test such as PISA, it is inevitable that the range of accents has to be restricted in order to ensure fairness to students of different abilities and with differing levels of exposure to particular accents, in order to avoid additional cognitive demands associated with adjusting to a new accent (Field, 2013[40]; Kang, Thomson and Moran, 2018[46]). Based on these considerations, the PISA test of English listening includes only British and North American varieties (avoiding all localisms or regional variations), in approximately equal proportions across all levels.

Overview of test tasks, levels and processes

Each assessment unit is designed to elicit evidence about a particular cognitive process involved in listening by adjusting the task characteristics described in the CEFR (e.g. listening situations, topics, speech rate, accent, etc.). At lower levels of proficiency, tasks assess a narrower range of processes involved in the literal understanding of what is being said. The simplest tasks focus on input decoding and lexical search and are based on short statements or short informal dialogues built around a single concept ("It is eleven o'clock", "What did you buy? I bought apples and oranges"). Input decoding tasks focus on the most basic vocabulary, testing students' ability to identify concrete information such as times, prices, days of the weeks and numbers when listening. Lexical search tasks (which also require input decoding) include a wider range of high-frequency vocabulary. Somewhat longer clauses are used to assess syntactic parsing, or the capacity to impose a syntactic structure onto words and understand the literal meaning of spoken clauses. More complex parsing tasks require listeners to identify key information or clearly stated opinions and may be built around longer texts that involve multiple speakers (e.g. in an interview). Stimuli are selected based on authentic scenarios (such as ordering food at a restaurant or a public announcement at a train station). Higher-level listening processes, involved in meaning construction and discourse construction are typically built around longer dialogues involving up to three speakers expressing opinions or feelings.

Table 4.6 provides an overview of the CEFR reception scales targeted at each level in the PISA Foreign Language Assessment listening test through a range of task types. The CEFR describes listening as taking place in different situations or contexts. The listening test for FLA is primarily concerned with situations involving overhearing a conversation between others, listening as part of an audience (seminar, tour) or listening to audio broadcasts.. These situations or contexts are considered appropriate and accessible to the age group of the PISA test takers.

Table 4.6. Overview of CEFR scales by proficiency level for listening scripts

CEFR level	CEFR scales represented
Pre-A1	Overall oral comprehension Identifying cues and inferring
A1	Overall oral comprehension Identifying cues and inferring Understanding conversation between other people Understanding announcements and instructions
A2	Overall oral comprehension Identifying cues and inferring Understanding conversation between other people Understanding announcements and instructions
B1	Overall oral comprehension Identifying cues and inferring Understanding conversation between other people Understanding audio media and recordings
B2	Overall oral comprehension Identifying cues and inferring Understanding conversation between other people Understanding audio media and recordings
C1	Overall oral comprehension Identifying cues and inferring Understanding conversation between other people Understanding audio media and recordings

Source: Cambridge Assessment English's original work, based on Council of Europe (2020[32]), *Common European Framework of Reference for Languages: Learning, teaching, assessment-Companion volume*, © Council of Europe, Strasbourg, pp. 48-53.

The listening test consists of a number of discrete assessment tasks, with some tasks corresponding to stand-alone items and longer tasks corresponding to multiple items sharing a common acoustic input. Although similar task types are used across different levels, their difficulty is differentiated by features of the listening script (described in the following section), by the testing focus, which increases in cognitive difficulty (e.g. identifying explicit detail, recognising speaker agreement, etc.), which are all dimensions described in the CEFR listening scales. Detailed mapping is provided in Annex 4.C. This mapping brings together task type, CEFR level, task description, cognitive processes and CEFR can-do statements. An extract from this mapping is presented in Table 4.7. Examples of the actual tasks can be found in Annex 4.D.

Table 4.8 provides approximate distributions of tasks by the cognitive process involved and target proficiency level.

Table 4.7. Extract from the mapping of listening tasks to CEFR level, cognitive process and CEFR can-do statements

Task type	CEFR level	Task description	Cognitive process	CEFR can-do statements
Listening: Discrete multiple-choice with graphics	Pre-A1	**Audio**: Short informal or neutral dialogue **Response format**: Question and three picture options **Operation**: Select the correct picture according to the audio	Input decoding	**Overall oral comprehension** Can understand short, very simple questions and statements, provided they are delivered slowly and clearly and accompanied by visuals or manual gestures to support understanding, and repeated if necessary. Can recognise everyday, familiar words/signs, provided they are delivered clearly and slowly in a clearly defined, familiar everyday context. Can recognise numbers, prices, dates and days of the week, provided they are delivered slowly and clearly in a defined, familiar everyday context. **Identifying cues and inferring** Can deduce the meaning of a word/sign from an accompanying picture or icon.

Note: See Annex 4.C. for the complete mapping table.
Source: Cambridge Assessment English's original work, based on Council of Europe (2020[32]), *Common European Framework of Reference for Languages: Learning, teaching, assessment – Companion volume*, © Council of Europe, Strasbourg, pp. 48-60 and Field (2013[40]), *Cognitive validity*, in Taylor, L. and A. Geranpayeh (eds.), *Examining Listening: Research and Practice in Assessing Second Language Listening,* UCLES/ Cambridge University Press, Cambridge, pp. 95-103..

Table 4.8. Distribution of listening tasks by process and proficiency level

	Input decoding	Lexical search	Parsing	Meaning construction	Discourse construction	Overall
Low (A1 and below)	50%	50%				30-35%
Medium (A2, B1)			100%			30-35%
High (B2 and above)				60%	40%	30-35%

Note: Lower-level cognitive processes (to the left) are nested within higher-order processes (to the right).
Source: Cambridge Assessment English's original work, based on Field, J. (2013[40]), *Cognitive validity*, in Taylor, L. and A. Geranpayeh (eds.), *Examining Listening: Research and Practice in Assessing Second Language Listening*, UCLES/Cambridge University Press, Cambridge, pp. 95-103.

Task presentation and response formats: Listening

In all listening tasks, aspects of context – such as test navigation interface and response formats – are kept constant, so as to avoid competing cognitive demands during the assessment. The channel of presentation is exclusively aural input (no video-recorded or visual input is used). Although the use of visual input is argued to support authenticity (Wagner and Ockey, 2018[47]), there is a danger that it can introduce construct irrelevant variance (Buck, 2001[48]). Additionally, research investigating the effect of visual input on listening behaviour and scores suggests that different types of visuals (context-setting visuals vs. content visuals; static vs. dynamic visuals) can have different effects on learners and scores (Suvorov, 2013[49]). This suggests it would not be appropriate at this time to include visual input.

When students start the listening test, they see the questions before the auditory input starts, and questions remain visible on the screen during the entire replay time (and until the student clicks on "next", confirming submission of his/her answer). Where multiple questions are asked about the same input, these are ordered in the same sequence as the relevant passages in the recording. In this way, competing cognitive demands are minimised. Students have the opportunity to listen to auditory stimulus twice, with a short pause before the second play; students cannot, however, pause the playback while answering questions, or sample a particular passage and listen to it again. While repetition is often necessary at lower levels of proficiency to ensure understanding, it may appear somewhat unnatural at higher levels of proficiency. However, Holzknecht (2019[50]) found that allowing audio to be played twice improved construct representation and use of higher-order cognitive processes over a single-play option. The response format used in PISA (selected response, with multiple distractors) further justifies the repetition of the auditory stimulus, to allow students to match their understanding with the written or visual answers provided, and check their answers.

As described in the reading section, the response format can support the activation of the cognitive processes of interest, and it can also be a source of construct irrelevant variance. This is particularly important for listening as students' listening competence must be measured via another skill (Field, 2013[40]). In order to try to reduce method effects while also activating a wide range of cognitive processes, only a single response format – multiple-choice selection – is used in the PISA Foreign Language Assessment listening test. This ensures that students' success in the test is not confounded by their ability to understand the task requirements or instructions because they are likely familiar with the task type. Furthermore, and particularly at the lower levels of proficiency, graphical answer options (such as weather icons) or numbers are included; students are therefore able to select the correct answer without needing to know the written form of a word. More generally, from Level A2 and above, the reading level required to answer the listening items has been kept at least one CEFR level below the target level assessed by the listening task.

Listening test forms

Similar to reading, the item pool is assembled in multiple and partially overlapping forms, with each student taking only one form. As this corresponds to a natural sequence of processes when listening, all students will start with tasks assessing the lower-level processes of input decoding, lexical search, and parsing (with the first tasks combining all three processes). By virtue of the adaptive assignment of students to forms, students who do well on the initial parsing tasks are expected to advance to more complex meaning construction and discourse construction tasks, while students who do not do well on the initial tasks will be assessed mostly on the component skills of input decoding and lexical search. This will, in turn, ensure that students are assessed on listening processes, and contexts that are represented in the CEFR proficiency descriptors, at their level of proficiency.

Assessing speaking in a foreign language

The speaking test is designed to elicit evidence about the cognitive processes involved in speaking, as described in Chapter 3 and in the CEFR descriptors for spoken production corresponding to different ability levels. The test is computer delivered. The CEFR scales related to production are the focus of the speaking test because it is not possible to adequately represent spoken interaction in a computer delivered speaking test such as the PISA Foreign Language Assessment.

In contrast to the reading and listening tests, the speaking test is non-adaptive. All students are presented with the same tasks, and it is the characteristics of the spoken output that determine the CEFR level. Thisis because, as indicated in Chapter 3, most of the five cognitive processes (i.e. planning, grammatical, phonological and phonetic encoding and self-monitoring) are required in spoken production regardless of

CEFR level. However, depending on the proficiency level of students, they may face performance limitations in some areas (e.g. phonetic encoding), because effort is being expended in other areas (e.g. lexical retrieval) (Field, 2011[51]).

The mark scheme used to assess spoken performance differentiates student performance in terms of the control they have over the cognitive processes involved in speaking in accordance with CEFR descriptors at the different ability levels.

Task characteristics that affect speaking processes

Although individual speaking tasks are not aligned to specific CEFR levels, contextual parameters (such as task purpose, response formats, time constraints, prompt characteristics) are carefully selected to ensure that students at all CEFR levels (Pre-A1 to C1) are able to demonstrate their speaking ability.

Task content and topics are selected that are considered suitable for 15-year-olds in terms of their world knowledge and interests. The focus on authentic tasks provides students with a realistic reason for speaking, which "goes beyond a ritual display of knowledge for assessment" (Shaw and Weir, 2007, p. 71[52]).

As noted in Chapter 3, the purpose of the task or what the learner will be asked to do with the language (e.g. the functional requirements) has been found to activate the grammatical structures needed (Galaczi and ffrench, 2011[53]). Typically, the task purpose is indicated in the rubric with terms such as "describe", "compare", and "justify". A useful model to better understand task purpose from the writing literature, but equally pertinent to speaking, is Scardamalia and Bereiter's (1987[54]) distinction between knowledge telling and transforming. Knowledge telling relates to sequential ordering of information and requires no manipulation of input or information, whereas knowledge transformation captures the ability to re-organise, evaluate and re-invent information in complex ways. As such, students will progress from the less cognitively demanding knowledge telling to the more cognitively demanding knowledge transformation as their proficiency increases. The task purpose is often encoded in the discourse mode, which helps students identify the rhetorical expectations of the task (i.e. narration, description, evaluation, etc.). As students' proficiency improves, they move from being able to cope with factual, descriptive modes to more abstract opinion or evaluative modes (Brown et al., 1984[55]; Bygate, 1987[56]). Therefore, at the lower levels, students are primarily engaged in knowledge telling with the focus on factual descriptions or simple narrations. From Level B2 upwards, students can engage in knowledge transformation in that they can make an argument or provide a justification for an opinion. As a result, the picture prompts allow students to engage in either knowledge telling and/or knowledge transformation depending on their proficiency levels.

Overview of test tasks and CEFR scales represented in the speaking test

The spoken production scales represented in the speaking test involve the functions of narrating, describing and sustaining an argument. Additionally, a number of scales from *Communicative Language Competences* are relevant, as these scales map most clearly to the cognitive processes for speaking. Table 4.9 provides an overview of the speaking tasks, and examples of them can be found in Annex 4.F. The overall oral production and phonological control in speaking can-do statements are presented in Annex 4.E.

Table 4.9. Overview of speaking tasks and CEFR scales represented in the speaking test

Task	Description	CEFR scales represented
Read aloud	Individual sentences/questions are presented on the screen one at a time, and the student reads each one out.	Phonological control
Odd-one out	Four pictures of objects/actions are presented on the screen, and the student explains which one does not belong in the set.	Sustained monologue: Putting a case General linguistic range Phonological control Thematic development Coherence and cohesion Fluency
Long turn – picture	A picture of things/people/situations is presented on the screen, and two separate questions are asked about the picture (one descriptive and one speculative).	Sustained monologue: Describing experience Sustained monologue: Putting a case General linguistic range Phonological control Thematic development Coherence and cohesion Fluency
Storytelling	A set of pictures is presented on the screen representing a storyboard. Student narrates a story based on the pictures.	Sustained monologue: Describing experience General linguistic range Phonological control Thematic development Coherence and cohesion Fluency
Long turn – graphic	A diagram (e.g. graphic organiser) with ideas related to a topic is presented on the screen, and the student is asked to talk about the information, giving their opinion and justifying it.	Sustained monologue: Putting a case General linguistic range Phonological control Thematic development Coherence and cohesion Fluency

Source: Cambridge Assessment English's original work, based on Council of Europe (2020[32]), *Common European Framework of Reference for Languages: Learning, teaching, assessment – Companion volume*, © Council of Europe, Strasbourg, pp. 62-64 and pp. 129-142.

As indicated in Table 4.9, the speaking test consists of a number of discrete assessment tasks, with all but one task corresponding to stand-alone items, and one longer task corresponding to multiple items sharing a common visual input (as is usual in PISA assessments).

After students complete a short practice task by answering some basic personal questions to familiarise themselves with the computer format, they read sentences and questions aloud, presented on the screen one at a time. This task is designed to assess students' phonetic encoding and articulation. Input sentences are arranged by difficulty from Level Pre-A1 to C1 based on their grammatical features and vocabulary. Students demonstrate their level of control over the pronunciation of individual sounds, the extent to which they can use stress, rhythm and intonation to express meaning as defined in the CEFR scale, "Phonological control".

In the remaining four tasks, students are presented with visual input in the form of a picture(s) or diagram (s) and answer one or two questions related to the input, presented visually on the screen and aurally. For example, to explain which picture in a set of four does not belong in the set; provide an opinion or preference related to a picture; tell a story based on the images they see; and speak at length on the information given in a graphic. Each of these tasks requires students to sustain a monologue, engaging in

all cognitive processes involved in speaking; however, the cognitive processes employed will vary according to the proficiency level of the student. Each task includes a context and a purpose for speaking.

The tasks are designed so that students at the lowest proficiency levels can focus on lexical retrieval (i.e. phonological encoding) and phonetic encoding by naming things they see in the images. As proficiency increases, grammatical encoding will be evidenced, and features related to conceptualisation appear as they organise their utterances in order to achieve the task's goals. The tasks provide students at higher proficiency levels opportunities to provide elaborated responses that move beyond description because they are designed to encourage students to speculate, offer an opinion and justify that opinion.

Response formats: Speaking

In speaking assessments, response formats can be classified as controlled, semi-controlled or open-ended, depending on how much the expected output is predetermined or controlled by the input. The more control students have over the expected response, the more open the task is, as students have to generate their own ideas. Controlled tasks reduce the cognitive load on students by "providing scaffolding upon which to build language" (Wigglesworth, 2001, p. 203[57]), whereas open-ended tasks place more cognitive demands on students because the effectiveness of their response will be heavily dependent on micro- and macro-planning (Field, 2011[51]). Additionally, responses to open-ended tasks can vary markedly, leading to marking issues in terms of comparability of scores. Therefore, open-ended speaking tasks are avoided. The speaking test tasks are ordered along a continuum from controlled (read aloud task) to semi-controlled (remaining four tasks ordered from most control to least control) to ensure tasks that appear earlier in the test provide some scaffolding for students, and encourage responses (see sample tasks in Annex 4.F).

Planning time

An important consideration is whether planning time is included. Wigglesworth (1997[58]) found that as tasks become cognitively more abstract and complex, planning time may help support learners at the conceptualisation stage; however, for less cognitively demanding tasks, planning time may not have any effect on learner language output (Iwashita, McNamara and Elder, 2001[59]) and could actually be detrimental to fluency (Wigglesworth, 2001[57]). The implication is that tasks that have a clear, well-defined structure do not need planning time. As students are provided with content or ideas that provide a springboard for their answers, planning time has not been included in addition to the 5-10 seconds required to process instructions and look at visuals. A by-product of not including planning time is that it will make students' ability to self-monitor more apparent in the recording. With little time to plan, students' utterances will likely include false starts, reformulations or fillers to gain planning time (e.g. "you know"). The extent to which they are able to identify and fix errors while speaking is an indicator of proficiency level (Field, 2011[51]).

Speaking prompts and instructions

As much as possible, speaking prompts are selected that rely on visual stimuli to reduce reliance on reading or listening comprehension and to prevent students from using the language of the prompt to formulate a response. When written input is included, it is also presented aurally to students, and the vocabulary and grammatical structures used are written to be accessible to students with lower-level proficiency.

Students see instructions on the screen while they also hear the same instructions spoken. For aural input, a "neutral" accent is used. This neutral accent is characterised as a hybrid of standard British and American accents in terms of vowel production and intonation, making it difficult to locate it in either country, but is

perceived as a neutral or international form of English (Modiano, 1996[31]). The instructions, visual input and any associated questions remain visible on the screen during the entire task. British spelling conventions are used for all written input presented on the screen. A progression bar indicates speaking time. The amount of time given to speak is clearly indicated in the instructions and next to the progression bar, to allow students to manage their planning and speaking time.

Notes

1. In the FLA Framework, "task" is used to refer to any purposeful language activity that language users and learners engage with; this includes, in the context of the foreign language test, any assessment items or exercises (e.g. a set of questions about a text), which are also referred to as (assessment) tasks.

2. As discussed in Chapter 2, foreign-language learners perform communicative language activities (reception, production, interaction and mediation) by using reading, listening, speaking and writing skills. Reading and listening are involved in reception; speaking and writing in production; and all four skills can be involved in interaction and mediation. PISA 2025 will assess only comprehension (reading and listening) and spoken production. It is the intention of the PISA Programme to explore assessing the remaining modes of communication in future iterations of its Foreign Language Assessment.

3. See Crossley, Greenfield and McNamara (2008[60]) for an overview of readability indices; for a discussion on criterial features, see Hawkins and Filipović (2012[61]) and Green (2012[23]).

4. The CEFR indicates that learners at Level C1 can understand speech delivered "at a natural speed in varieties that are familiar", suggesting that comprehension of accents is very much dependent on exposure.

References

Alderson, J. (2000), *Assessing Reading*, Cambridge University Press, Cambridge. [13]

Alderson, J., C. Clapham and D. Wall (1995), *Language Test Construction and Evaluation*, Cambridge University Press, Cambridge. [33]

Alderson, J. et al. (2004), *The Development of Specifications for Item Development and Classification within The Common European Framework of Reference for Languages: Learning, Teaching, Assessment Reading and Listening Final Report of The Dutch CEF Construct Project*, Lancaster University, Lancaster, http://www.ling.lancs.ac.uk/cefgrid. [16]

Brown, G. et al. (1984), *Teaching Talk: Strategies for Production and Assessment*, Cambridge University Press, Cambridge. [55]

Brown, G. and G. Yule (1983), *Teaching the Spoken Language*, Cambridge University Press, Cambridge. [45]

Brown, J. and T. Hudson (1998), "The Alternatives in Language Assessment", *TESOL Quarterly*, Vol. 32/4, pp. 653-675, http://dx.doi.org/10.2307/3587999. [35]

Buck, G. (2001), *Assessing Listening*, Cambridge University Press, Cambridge. [48]

Bygate, M. (1987), *Speaking*, Oxford University Press, Oxford. [56]

Cambridge University Press (n.d.), *English Profile- what the CEFR means for English*, http://www.englishprofile.org/ (accessed on 18 June 2020). [27]

Chafe, W. (1979), "The flow of thought and the flow of language", in Givon, T. (ed.), *Syntax and Semantics, 12: Discourse and Syntax*, Academic Press, New York. [41]

Chiswick, B. and P. Miller (2005), "Linguistic distance: A quantitative measure of the distance between English and other languages", *Journal of Multilingual and Multicultural Development*, Vol. 26/1, pp. 1-11, http://dx.doi.org/10.1080/14790710508668395. [3]

Choi, Y. and D. Zhang (2018), "The Relative Role of Vocabulary and Grammatical Knowledge in L2 Reading Comprehension: A Systematic Review of Literature", *International Review of Applied Linguistics in Language Teaching*, https://doi.org/10.1515/iral-2017-0033. [25]

Corder S.P. (1981), "Language distance and the magnitude of the language learning task", in *Error Analysis and Interlanguage*, Oxford University Press, Oxford. [2]

Council of Europe (2020), *Common European Framework of Reference for Languages : learning, teaching, assessment. Companion volume.*, Council of Europe, Strasbourg, https://www.coe.int/en/web/education/-/common-european-framework-of-reference-for-languages-l. [32]

Council of Europe (2001), *Common European Framework of Reference for Languages. Learning, Teaching, Assessment*, Cambridge University Press, Cambridge, http://www.coe.int/lang-CEFR. [17]

Council of Europe (n.d.), *Reference Level Descriptions (RLDs) language by language*, https://www.coe.int/en/web/common-european-framework-reference-languages/reference-level-descriptions (accessed on 7 June 2020). [26]

Crossley, S., J. Greenfield and D. McNamara (2008), "Assessing text readability using cognitively based indices", *TESOL Quarterly*, Vol. 42/3, pp. 475-493, http://dx.doi.org/10.1002/j.1545-7249.2008.tb00142.x. [60]

Elder, C. and A. Davies (1998), "Performance on ESL examinations: Is there a language distance effect?", *Language and Education*, Vol. 12/1, pp. 1-17, http://dx.doi.org/10.1080/09500789808666736. [4]

Ellis, R. (2015), *Understanding second language acquisition*, Oxford University Press, Oxford. [43]

Field, J. (2013), "Cognitive validity", in Taylor, L. and A. Geranpayeh (eds.), *Examining Listening: Research and Practice in Assessing Second Language Listening*, UCLES/Cambridge University Press, Cambridge. [40]

Field, J. (2011), "Cognitive validity", in Taylor, L. (ed.), *Examining Speaking: Research and Practice in Assessing Second Language Speaking*, UCLES/Cambridge University Press, Cambridge. [51]

Foley, C. and S. Flynn (2013), "The role of the native language", in *The Cambridge Handbook of Second Language Acquisition*, Cambridge University Press, Cambridge, http://dx.doi.org/10.1017/CBO9781139051729.008. [7]

Galaczi, E. and A. ffrench (2011), "Context validity", in Taylor, L. (ed.), *Examining Speaking: Research and Practice in Assessing Second Language Speaking*, UCLES/Cambridge University Press, Cambridge. [53]

Gass S.M. and L. Selinker (1983), *Language Transfer in Language Learning*, Newbury House, Rowley, MA. [1]

Goldman-Eisler, F. (1968), *Psycholinguistics: Experiments in Spontaneous Speech*, Academic Press, London. [44]

Graves, M. et al. (1991), "Improving Instructional Texts: Some Lessons Learned", *Reading Research Quarterly*, Vol. 26/2, pp. 110-132. [34]

Green, A. (2012), "Language Functions Revisited: Theoretical and Empirical Bases for Language Construct Definition Across the Ability Range"*, English Profile Studies*, No. 2, Cambridge University Press, Cambridge. [23]

Hardison, C. et al. (2012), *Second-Language Skills for All? Analyzing a Proposed Language Requirement for US Air Force Officers*. [12]

Hawkins, J. and L. Filipović (2012), *Criterial features in L2 English: Specifying the reference levels of the Common European Framework*, Cambridge University Press. [61]

Holzknecht, F. (2019), *Double play in listening assessment (Doctoral Thesis)*, Lancaster University, https://eprints.lancs.ac.uk/id/eprint/139699/1/2019holzknechtphd.pdf (accessed on 21 September 2020). [50]

Hughes, A. (2003), *Testing for Language Teachers*, Cambridge University Press, Cambridge. [37]

Hu, M. and P. Nation (2000), "Unknown Vocabulary Density and Reading Comprehension", *Reading in a Foreign Language*, Vol. 13, pp. 403-430. [28]

Iwashita, N., T. McNamara and C. Elder (2001), "Can we predict task difficulty in an oral proficiency test? Exploring the potential of an information- processing approach to test design", *Language Learning*, Vol. 51/3, pp. 401-436. [59]

Jackobson, R. (1960), "Linguistics and Poetics", in Seboek, T. (ed.), *Style in Language*, John Wiley, New York. [14]

Kang, O., R. Thomson and M. Moran (2018), "Empirical Approaches to Measuring the Intelligibility of Different Varieties of English in Predicting Listener Comprehension of Tests", *Language Learning*, Vol. 68/1, pp. 115-146, http://dx.doi.org/10.1111/lang.12270. [46]

Khalifa, H. and C. Weir (2009), *Examining Reading: Research and Practice in Assessing Second Language Learning*, UCLES/Cambridge University Press, Cambridge. [24]

Koda, K. (2005), *Insights into Second Language Reading: A Cross-linguistic Approach*, Cambridge University Press, New York. [8]

Lardiere, D. (2009), "Some thoughts on the contrastive analysis of features in second language acquisition", pp. 173-227. [6]

Lindgren, E. and C. Muñoz (2013), "The influence of exposure, parents, and linguistic distance on young European learners' foreign language comprehension", *International Journal of Multilingualism*, Vol. 10/1, pp. 105-129, http://dx.doi.org/10.1080/14790718.2012.679275. [10]

Modiano, M. (1996), "The Americanization of Euro-English", *World Englishes*, Vol. 15/2, pp. 207-215. [31]

Nation, P. (2006), "How large a vocabulary is needed for reading and listening?", *Canadian Modern Language Review*, Vol. 63/1, pp. 59-82, http://dx.doi.org/10.3138/cmlr.63.1.59. [29]

Odlin, T. (1989), *Language Transfer: Cross-Linguistic Influence in Language Learning*, Cambridge University Press, http://dx.doi.org/10.1017/CBO9781139524537. [5]

Ringbom, H. (1987), *The Role of the First Language in Foreign Language Learning*, Multilingual Matters, Clevedon. [11]

Rost, M. (2002), *Teaching and Researching Listening*, Pearson Education, Harlow. [9]

Scardamalia, M. and C. Bereiter (1987), "Knowledge telling and knowledge transforming in written composition", in Rosenberg, S. (ed.), *Advances in Applied Psycholinguistics Volume 2: Reading, writing and language learning*, Cambridge University Press, Cambridge. [54]

Schmitt, N., X. Jiang and W. Grabe (2011), "The Percentage of Words Known in a Text and Reading Comprehension", *Modern Language Journal*, Vol. 95/1, pp. 26-43, http://dx.doi.org/10.1111/j.1540-4781.2011.01146.x. [30]

Schwanenflugel, P. and R. Stowe (1989), "Context Availability and the Processing of Abstract and Concrete Words in Sentences", *Reading Research Quarterly*, Vol. 24/1, pp. 114-126. [42]

Shaw, S. and C. Weir (2007), *Examining Writing: Research and Practice in Assessing Second Language Writing*, UCLES/Cambridge University Press, Cambridge. [52]

Suvorov, R. (2013), *Interacting with visuals in L2 listening tests: An eye-tracking study (Doctoral thesis)*, Iowa State University. [49]

Urquhart, A. (1984), "The Effect of Rhetorical Ordering on Readability", in Alderson, J. and A. Urquhart (eds.), *Reading in a Foreign Language*, Longman, London. [15]

Urquhart, A. and C. Weir (1998), *Reading in a Second Language: Process, Product and Practice*, Longman, New York. [36]

Van Ek, J. and J. Trim (2001), *Vantage*, Cambridge University Press, Cambridge. [20]

Van Ek, J. and J. Trim (1998), *Threshold 1990*, Cambridge University Press, Cambridge. [18]

Van Ek, J. and J. Trim (1998), *Waystage 1990*, Cambridge University Press,. [19]

Wagner, E. (2013), "Assessing listening", in *The companion to language assessment.*, John Wiley, Hoboken, NJ. [39]

Wagner, E. and G. Ockey (2018), "An overview of the use of audio-visual texts on L2 listening tests", in *Assessing L2 listening: Moving towards authenticity.*, John Benjamins, Amsterdam. [47]

Weir, C. (1993), *Understanding and Developing Language Tests*, Prentice Hall, New York. [38]

Wigglesworth, G. (2001), "Influences on Performance in Task-based Oral Assessments", in Bygate, M., P. Skehan and M. Swain (eds.), *Researching Pedagogic Tasks*, Longman, Harlow. [57]

Wigglesworth, G. (1997), "An investigation of planning time and proficiency level on oral test discourse", *Language Testing* 14, pp. 85-106. [58]

Wilkins, D. (1976), *Notional Syllabuses*, Oxford University Press, Oxford. [22]

Wilkins, D. (1973), “The Linguistic and Situational Content of the Common Core in a Unit/Credit System”, in Trim, J. et al. (eds.), *Systems Development in Adult Language Learning*, Council of Europe, Strasbourg. [21]

Annex 4.A. Mapping of reading tasks to framework dimensions

Annex Table 4.A.1 maps the PISA Foreign Language Assessment reading tasks to the relevant Common European Framework of Reference for Languages (CEFR) can-do statements. In addition to the CEFR mapping, the table captures task- and CEFR level information in relation to the cognitive processes involved, task description and reading type. Points to note are as follows:

- The *Overall reading comprehension* descriptor is included for each task and level to act as a reminder of the broad reading ability in action at each level. Certain other scales and descriptors frequently appear across the table, e.g. the scale *Identifying cues and inferring.* These frequently-represented scales, and the selected descriptors from these scales, highlight those abilities employed widely across a range of reading task types. An example of a descriptor from the above-mentioned scale at Level A1 is *Can deduce the meaning of an unknown word for a concrete action or object, provided the surrounding text is very simple, and on a familiar everyday subject.* This can-do statement will have general applicability to a number of different task types and, as such, appears in the table in several places.
- Some scales feature less frequently across the table due to their more specific relevance to certain task types, e.g. *Reading for orientation.*
- The featured descriptors from the relevant scales highlight the particular abilities used in specific tasks at certain levels for the successful completion of that reading task.
- Some scales are not represented on the table of reading tasks (e.g. *Reading for leisure*) as these are not currently part of the test construct.
- Descriptors from plus-levels are not included in the table.
- The highest level of cognitive processing required for each task is included, with the assumption that the other lower-level process will also be utilised.
- Please note that aspects of the can-do statements that are not relevant to the PISA cohort have been struck through.

Annex Table 4.A.1. Mapping of reading tasks to CEFR level, cognitive process, CEFR can-do statements and reading type

Task type	CEFR level	Task description	Cognitive process	CEFR can-do statements	Reading type
Reading – Discrete definition with graphic	Pre-A1	**Input text**: Picture **Response format**: Three short sentences **Operation**: Select the sentence that defines the picture	Word recognition and lexical access	**Overall reading comprehension** Can recognise familiar words/signs accompanied by pictures, such as a fast-food restaurant menu illustrated with photos or a picture book using familiar vocabulary. **Identifying cues and inferring** Can deduce the meaning of a word from an accompanying picture or icon.	Careful, local
Reading – Matching pictures and words	Pre-A1	**Input text**: Five pictures **Response format**: Seven words **Operation**: Match the word to the correct picture – two distractors	Word recognition and lexical access	**Overall reading comprehension** Can recognise familiar words accompanied by pictures, such as a fast-food restaurant menu illustrated with photos or a picture book using familiar vocabulary. **Identifying cues and inferring** Can deduce the meaning of a word from an accompanying picture or icon.	Careful, local
Reading – Matching sentences	A1	**Input text**: Five sentences **Response format**: Seven words **Operation**: Match the word to the correct sentence – two distractors	Establishing propositional meaning at sentence level	**Overall reading comprehension** Can understand very short, simple texts a single phrase at a time, picking up familiar names, words and basic phrases and rereading as required. **Identifying cues and inferring** Can deduce the meaning of an unknown word for a concrete action or object, provided the surrounding text is very simple, and on a familiar everyday subject.	Careful, local
Reading – Discrete gapped sentence	A1	**Input text**: Sentence with a word missing (gap) **Response format**: Three words **Operation**: Select the word that fills the gap to complete the sentence	Lexical access/propositional meaning at sentence level	**Overall reading comprehension** Can understand very short, simple texts a single phrase at a time, picking up familiar names, words and basic phrases and rereading as required. **Identifying cues and inferring** Can deduce the meaning of an unknown word for a concrete action or object, provided the surrounding text is very simple, and on a familiar everyday subject.	Careful, local
	A2	**Input text**: Sentence with a word missing (gap) **Response format**: Three words **Operation**: Select the word that fills the gap to complete the sentence	Lexical access/propositional meaning at sentence level	**Overall reading comprehension** Can understand short, simple texts containing the highest frequency vocabulary, including a proportion of shared international vocabulary items.	Careful, local
Reading – Discrete multiple-choice	A1	**Input text**: Short text (notice, sign, message) **Response format**: Three sentences	Propositional meaning at sentence level	**Overall reading comprehension** Can understand very short, simple texts a single phrase at a time, picking up familiar names, words and basic phrases and rereading as required.	Careful, local

Task type	CEFR level	Task description	Cognitive process	CEFR can-do statements	Reading type
		Operation: Select the sentence that is true according to the text		**Reading correspondence** Can understand short, simple messages on postcards. Can understand short, simple messages sent via social media or email (e.g. proposing what to do, when and where to meet). **Identifying cues and inferring** Can deduce the meaning of an unknown word for a concrete action or object, provided the surrounding text is very simple, and on a familiar everyday subject.	
	A2	**Input text**: Short text (notice, sign, message) **Response format**: Three sentences **Operation**: Select the sentence that is true according to the text	Propositional meaning at sentence level	**Overall reading comprehension** Can understand short, simple texts containing the highest frequency vocabulary, including a proportion of shared international vocabulary items. **Reading correspondence** Can understand short, simple personal letters. Can understand very simple formal emails and letters (e.g. confirmation of a booking or online purchase). **Identifying cues and inferring** Can exploit format, appearance and typographic features in order to identify the type of text: news story, promotional text, article, textbook, chat or forum, etc. Can exploit numbers, dates, names, proper nouns, etc. to identify the topic of a text. Can deduce the meaning and function of unknown formulaic expressions from their position in a text (e.g. at the beginning or end of a letter).	Careful, local
	B1	**Input text**: Short text (notice, sign, message) **Response format**: Three sentences **Operation**: Select the sentence that is true according to the text	Inferencing	**Overall reading comprehension** Can read straightforward factual texts on subjects related to their field of interest with a satisfactory level of comprehension. **Reading correspondence** Can understand the description of events, feelings and wishes in personal letters well enough to correspond regularly with a pen pal. Can understand straightforward personal letters, emails or postings giving a relatively detailed account of events and experiences. **Identifying cues and inferring** Can make basic inferences or predictions about text content from headings, titles or headlines. Can deduce the probable meaning of unknown words/signs in a text by identifying their constituent parts (e.g. identifying roots, lexical elements, suffixes and prefixes).	Careful, local and global
Reading – Multi-item multiple-choice	B1	**Input text**: Long text	Building a mental model	**Overall reading comprehension** Can read straightforward factual texts on subjects related to their field of interest with a satisfactory level of comprehension.	Careful, local and global

Task type	CEFR level	Task description	Cognitive process	CEFR can-do statements	Reading type
		Response format: Five questions with four options per question **Operation**: Select the option for each question that is true according to the text		**Reading for information and argument** Can recognise significant points in straightforward news articles on familiar subjects. Can understand most factual information that they are likely to come across on familiar subjects of interest, provided they have sufficient time for rereading. **Identifying cues and inferring** Can make basic inferences or predictions about text content from headings, titles or headlines. Can follow a line of argumentation or the sequence of events in a story, by focusing on common logical connectors (e.g. however, because) and temporal connectors (e.g. after that, beforehand). Can deduce the probable meaning of unknown words/signs in a text by identifying their constituent parts (e.g. identifying roots, lexical elements, suffixes and prefixes).	
	B2	**Input text**: Long text **Response format**: 1) Five questions with four options per question 2) Two questions with four options per question **Operation**: Select the option for each question that is true according to the text	Building a mental model	**Overall reading comprehension** Can read with a large degree of independence, adapting style and speed of reading to different texts and purposes, and using appropriate reference sources selectively. Has a broad active reading vocabulary, but may experience some difficulty with low-frequency idioms. **Reading for information and argument** Can understand articles and reports concerned with contemporary problems in which particular stances or viewpoints are adopted. Can recognise when a text provides factual information and when it seeks to convince readers of something. Can recognise different structures in discursive text: contrasting arguments, problem–solution presentation and cause–effect relationships. **Identifying cues and inferring** Can use a variety of strategies to achieve comprehension, including watching out for main points and checking comprehension by using contextual clues.	Careful, local and global
	C1	**Input text**: Long text **Response format**: Five questions with four options per question **Operation**: Select the option for each question that is true according to the text	Building a mental model	**Overall reading comprehension** Can understand in detail lengthy, complex texts, whether or not these relate to their own area of speciality, provided they can reread difficult sections. Can understand a wide variety of texts, including literary writings, newspaper or magazine articles, and ~~specialised academic or professional~~ publications, provided there are opportunities for rereading ~~and they have access to reference tools~~. **Reading for information and argument**	Careful, local and global

Task type	CEFR level	Task description	Cognitive process	CEFR can-do statements	Reading type
				Can understand in detail a wide range of lengthy, complex texts likely to be encountered in social, ~~professional~~ or academic life, identifying finer points of detail including attitudes, and implied as well as stated opinions. **Identifying cues and inferring** Is skilled at using contextual, grammatical and lexical cues to infer attitude, mood and intentions, and anticipate what will come next.	
Reading – Matching gapped text – sentences	B1	**Input text**: Long text with five gaps representing missing sentences **Response format**: Eight sentences **Operation**: Match the missing sentence to the correct gap so that the text is coherent	Building a mental model	**Overall reading comprehension** Can read straightforward factual texts on subjects related to their field of interest with a satisfactory level of comprehension. **Reading for information and argument** Can recognise significant points in straightforward news articles on familiar subjects. Can understand most factual information that they are likely to come across on familiar subjects of interest, provided they have sufficient time for rereading. **Identifying cues and inferring** Can make basic inferences or predictions about text content from headings, titles or headlines. Can follow a line of argumentation or the sequence of events in a story, by focusing on common logical connectors (e.g. however, because) and temporal connectors (e.g. after that, beforehand). Can deduce the probable meaning of unknown words in a text by identifying their constituent parts (e.g. identifying roots, lexical elements, suffixes and prefixes).	Careful, global
Reading – Matching gapped text – paragraphs	C1	**Input text**: Long text with five gaps representing missing paragraphs **Response format**: Seven paragraphs **Operation**: Match the missing paragraph to the correct gap so that the text is coherent	Creating a text-level structure	**Overall reading comprehension** Can understand in detail lengthy, complex texts, whether or not these relate to their own area of speciality, provided they can reread difficult sections. Can understand a wide variety of texts including literary writings, newspaper or magazine articles, and ~~specialised~~ academic ~~or professional~~ publications, provided there are opportunities for rereading ~~and they have access to reference tools.~~ **Reading for information and argument** Can understand in detail a wide range of lengthy, complex texts likely to be encountered in social,~~ professional~~ or academic life, identifying finer points of detail, including attitudes and implied, as well as stated opinions. **Identifying cues and inferring** Is skilled at using contextual, grammatical and lexical cues to infer attitude, mood and intentions, and anticipate what will come next.	Careful, global
Reading – Multiple matching	B2	**Input text**: Four short texts on a related theme	Building a mental model	**Overall reading comprehension**	Expeditious, local – search reading

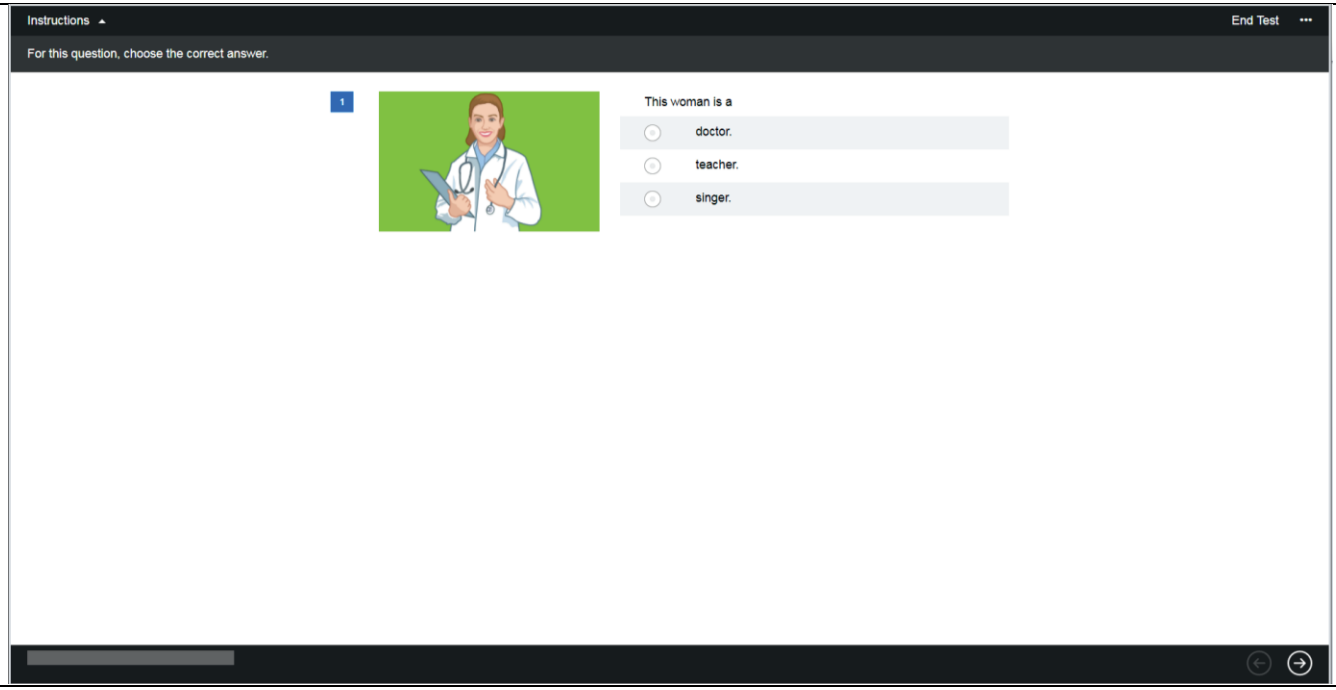
Task type	CEFR level	Task description	Cognitive process	CEFR can-do statements	Reading type
		Response format: Ten questions **Operation**: Select the text that answers the question (Note: There is a time limit of five minutes to test expeditious reading).		Can read with a large degree of independence, adapting style and speed of reading to different texts and purposes, and using appropriate reference sources selectively. Has a broad active reading vocabulary, but may experience some difficulty with low-frequency idioms. **Reading for orientation** Can scan quickly through long and complex texts, locating relevant details. Can quickly identify the content and relevance of news items, articles and reports on a wide range of professional topics, deciding whether closer study is worthwhile. **Reading for information and argument** Can understand articles and reports concerned with contemporary problems in which particular stances or viewpoints are adopted. Can recognise different structures in discursive text: contrasting arguments, problem–solution presentation and cause–effect relationships. **Identifying cues and inferring** Can use a variety of strategies to achieve comprehension, including watching out for main points and checking comprehension by using contextual clues.	
Reading – Cross-text matching	C1	**Input text**: Four short texts on a related theme **Response format**: Four questions **Operation**: Select the text that answers the question	Creating an intertextual representation	**Overall reading comprehension** Can understand in detail lengthy, complex texts, whether or not these relate to their own area of speciality, provided they can reread difficult sections. Can understand a wide variety of texts including literary writings, newspaper or magazine articles, and ~~specialised~~ academic ~~or professional~~ publications, provided there are opportunities for rereading ~~and they have access to reference tools.~~ **Reading for orientation** Can scan quickly through several sources (articles, reports, websites, books, etc.) in parallel, in both their own field and in related fields, and can identify the relevance and usefulness of particular sections for the task at hand. **Reading for information and argument** Can understand in detail a wide range of lengthy, complex texts likely to be encountered in social, ~~professional~~ or academic life, identifying finer points of detail, including attitudes, and implied as well as stated opinions. **Identifying cues and inferring** Is skilled at using contextual, grammatical and lexical cues to infer attitude, mood and intentions, and anticipate what will come next.	Careful, global

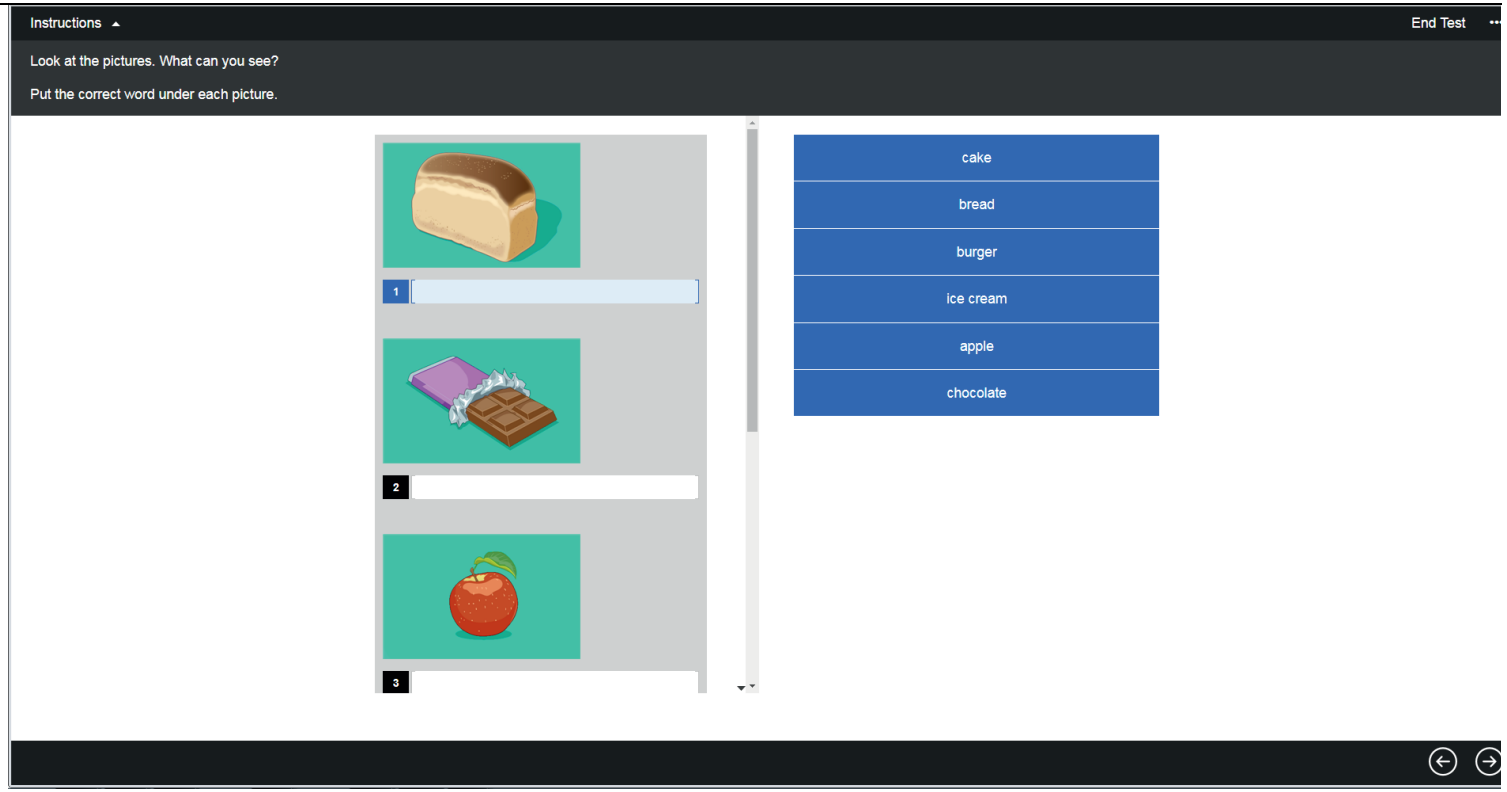
Source: Cambridge Assessment English's original work, based on Khalifa, H. and C.J. Weir (2009[24]), *Examining Reading: Research and Practice in Assessing Second Language Learning*, UCLES/Cambridge University Press, Cambridge, p. 70 and Council of Europe (2020[32]), *Common European Framework of Reference for Languages: Learning, teaching, assessment – Companion volume*, © Council of Europe, Strasbourg, pp. 54-60.

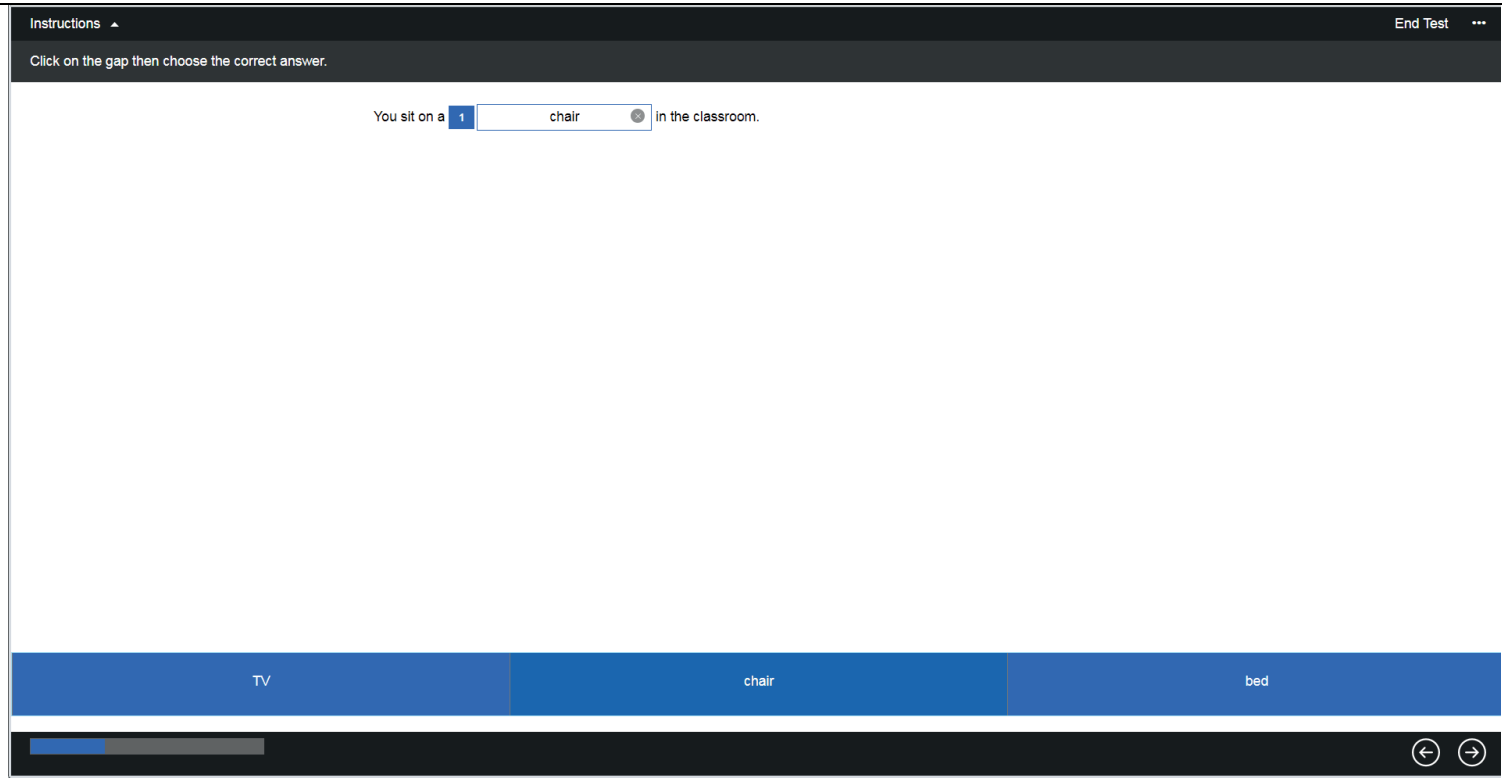
Annex 4.B. Sample tasks: Reading

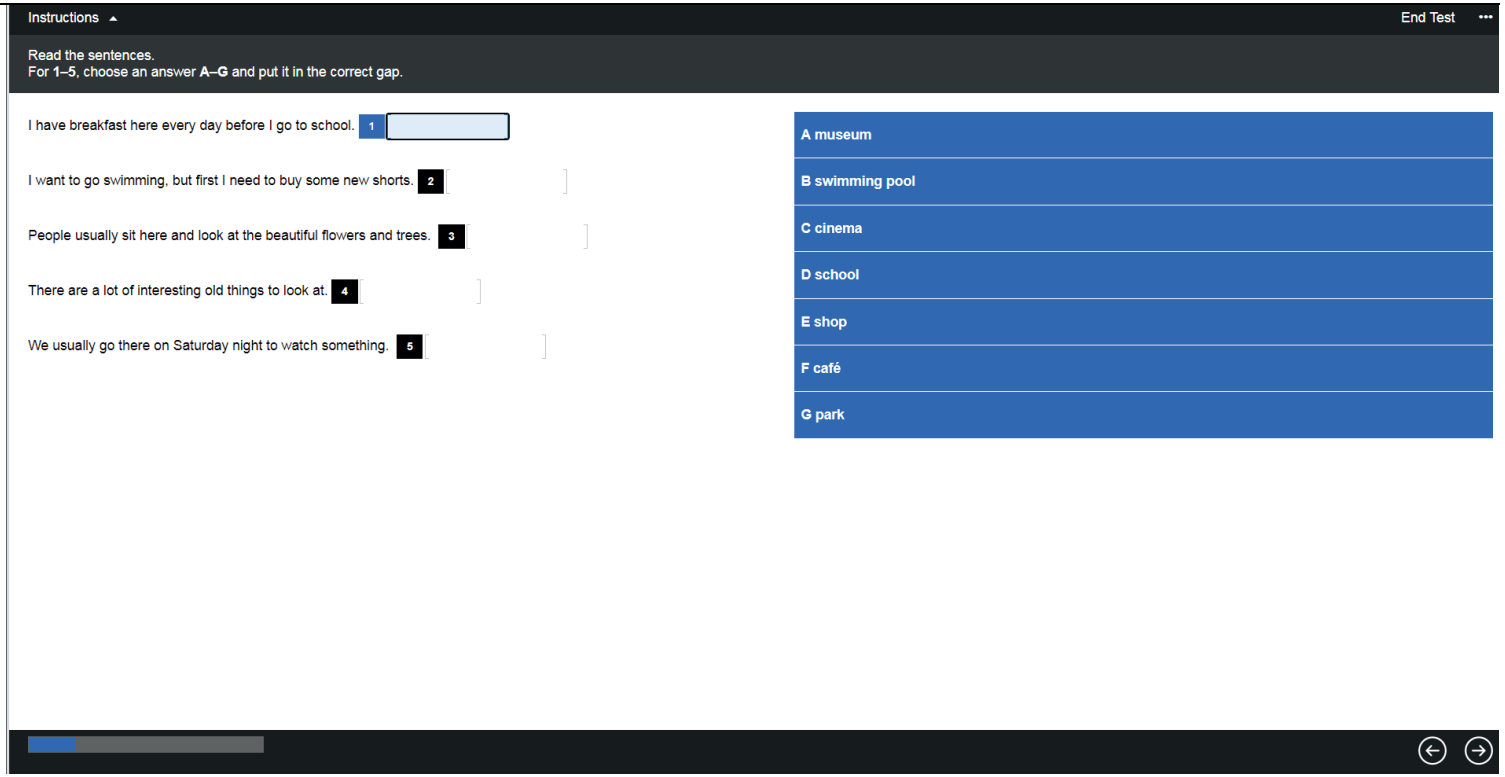
An example is given for each task type. Please note the same task type may be used at a number of different levels. Where a task type is used at more than one level, the level of the example is shown in bold.

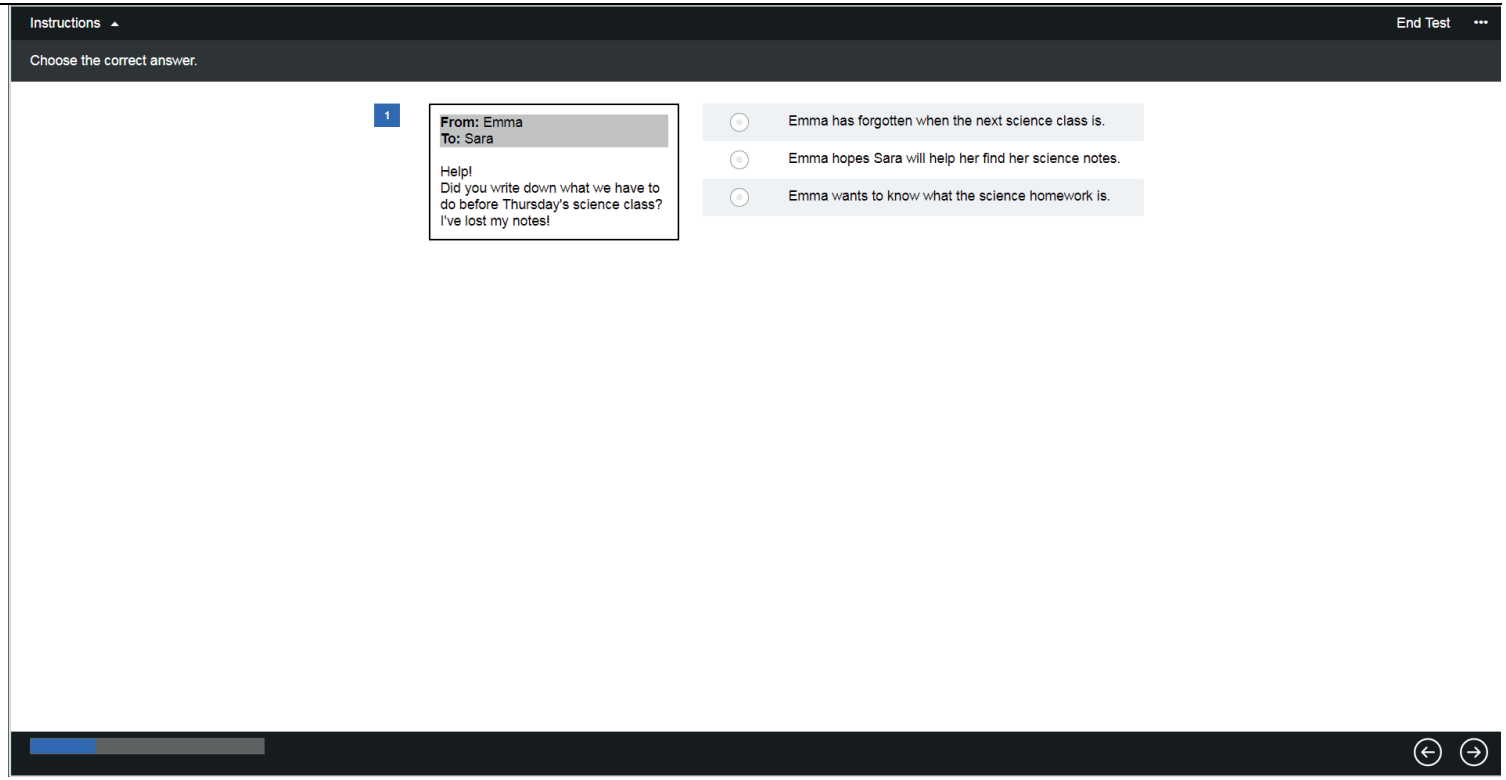
Annex Table 4.B.1. Sample reading tasks

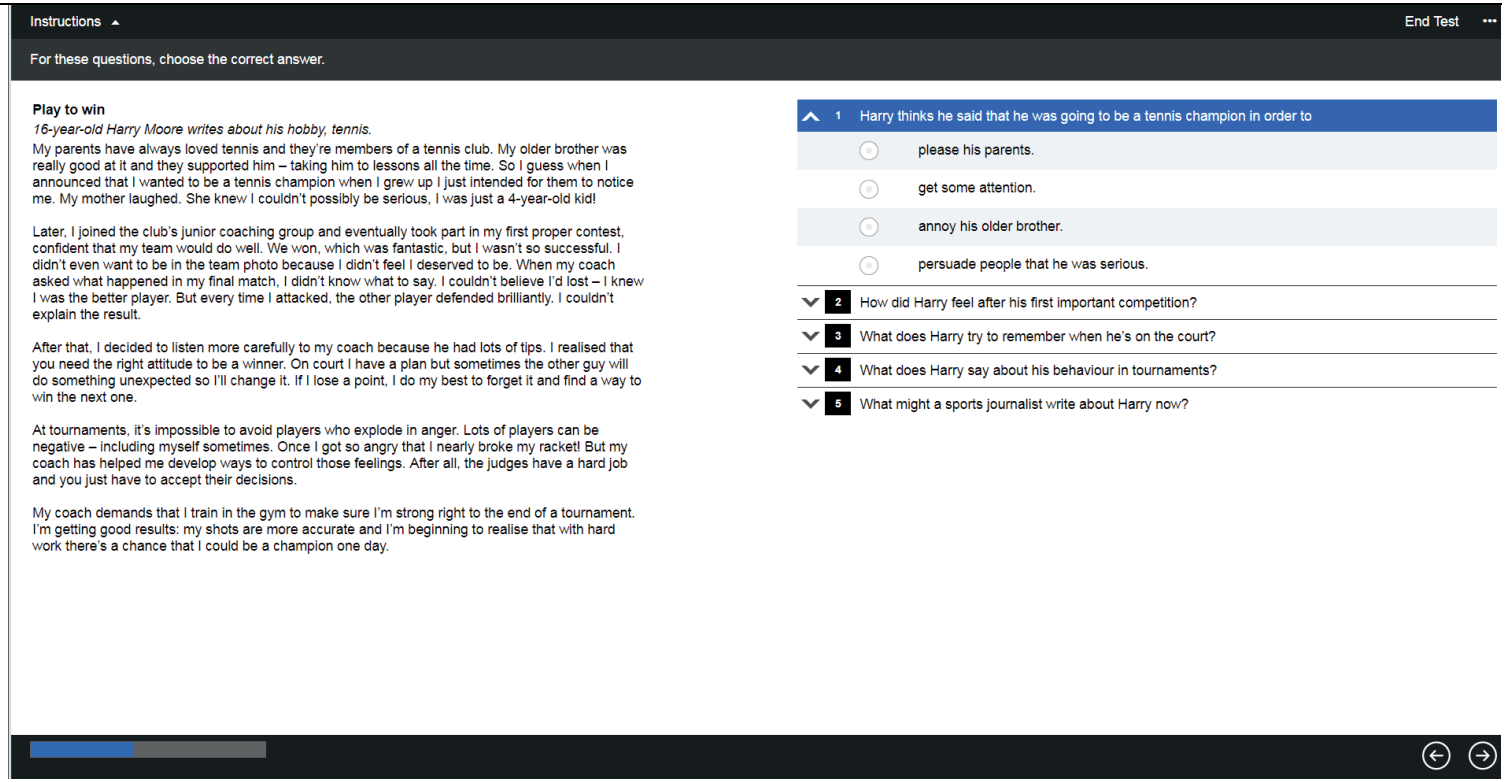
Task type	CEFR level	Sample task
Discrete definition with graphic	**Pre-A1**	

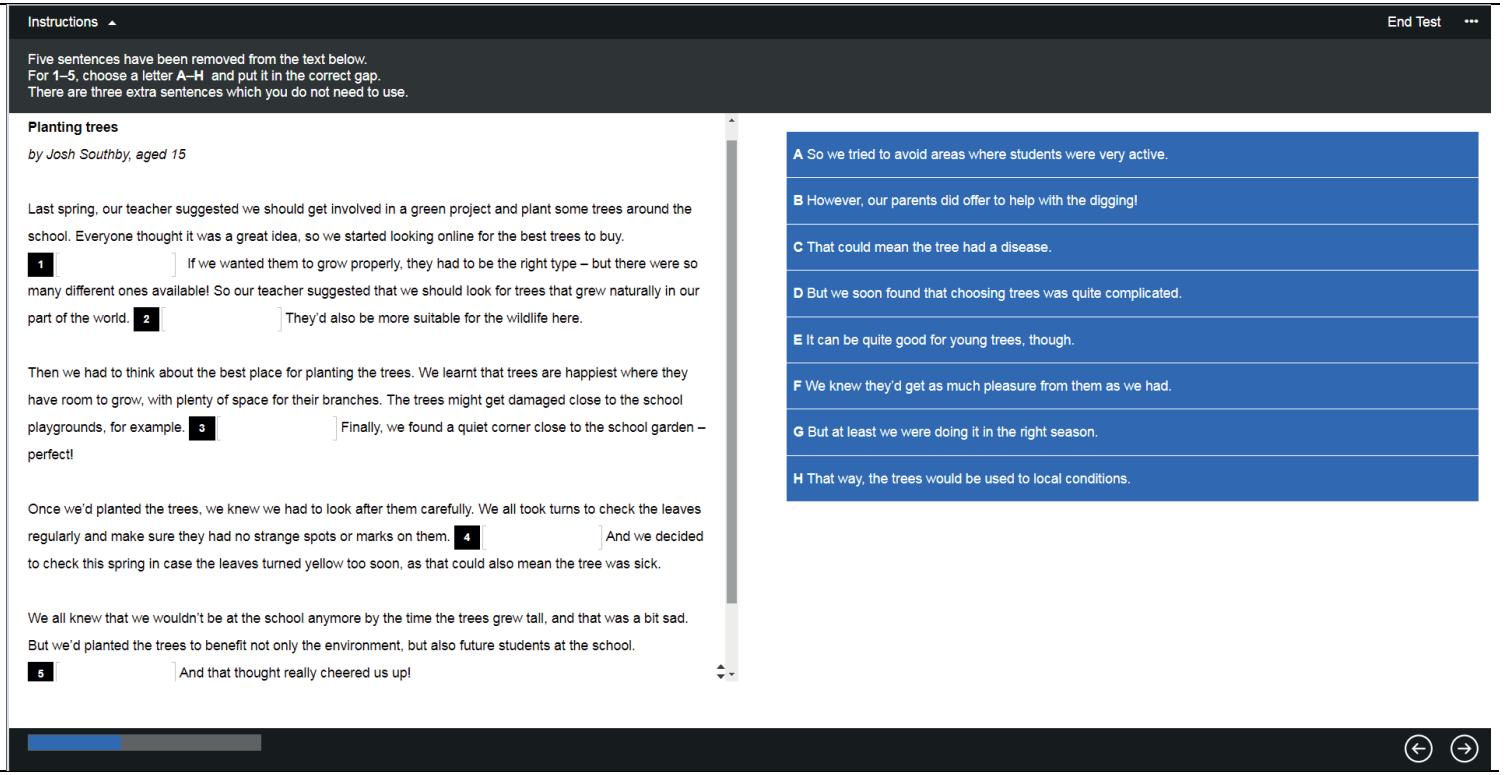
Task type	CEFR level	Sample task
Matching picture and words	**Pre-A1**	

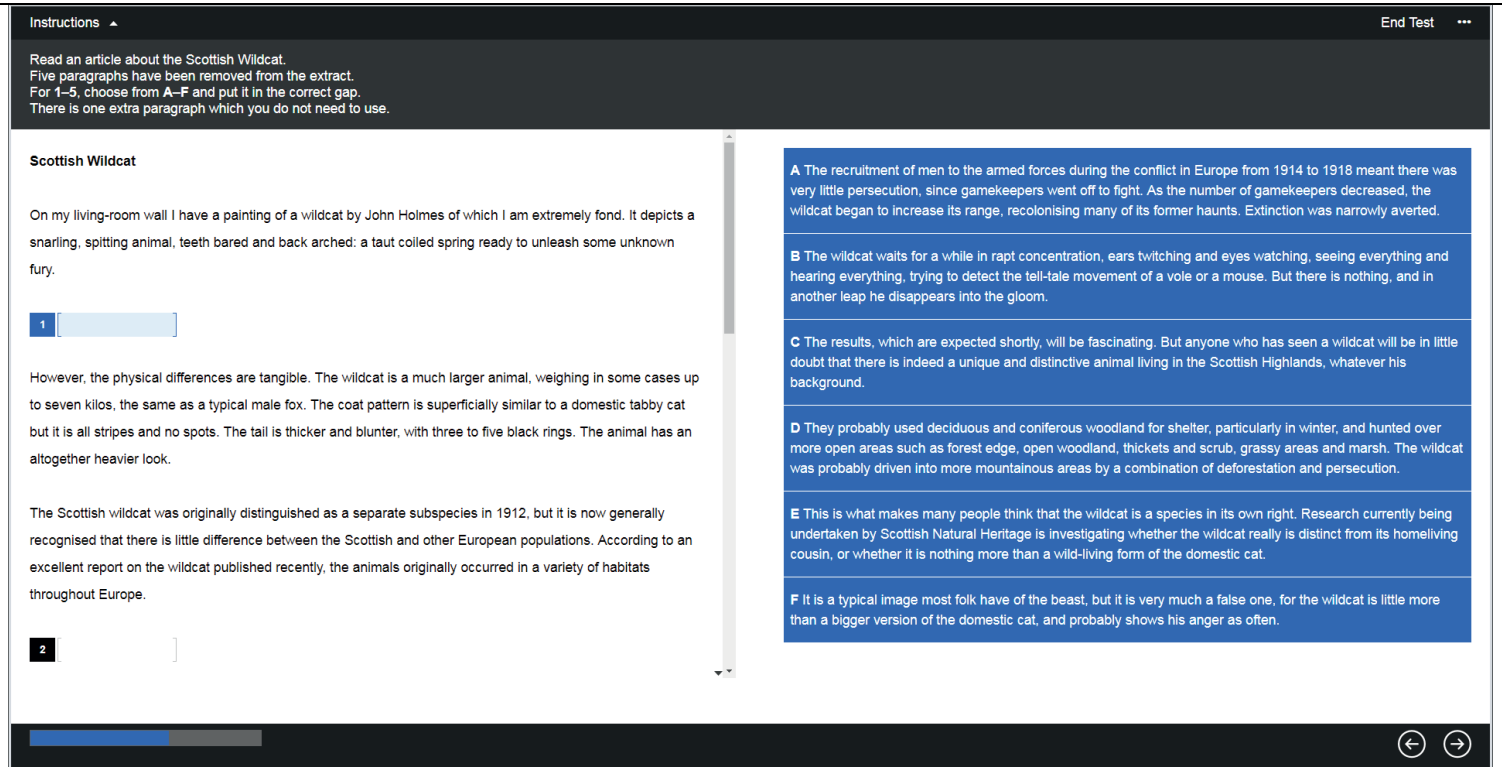
Task type	CEFR level	Sample task
Discrete gapped sentence	**A1** A2	

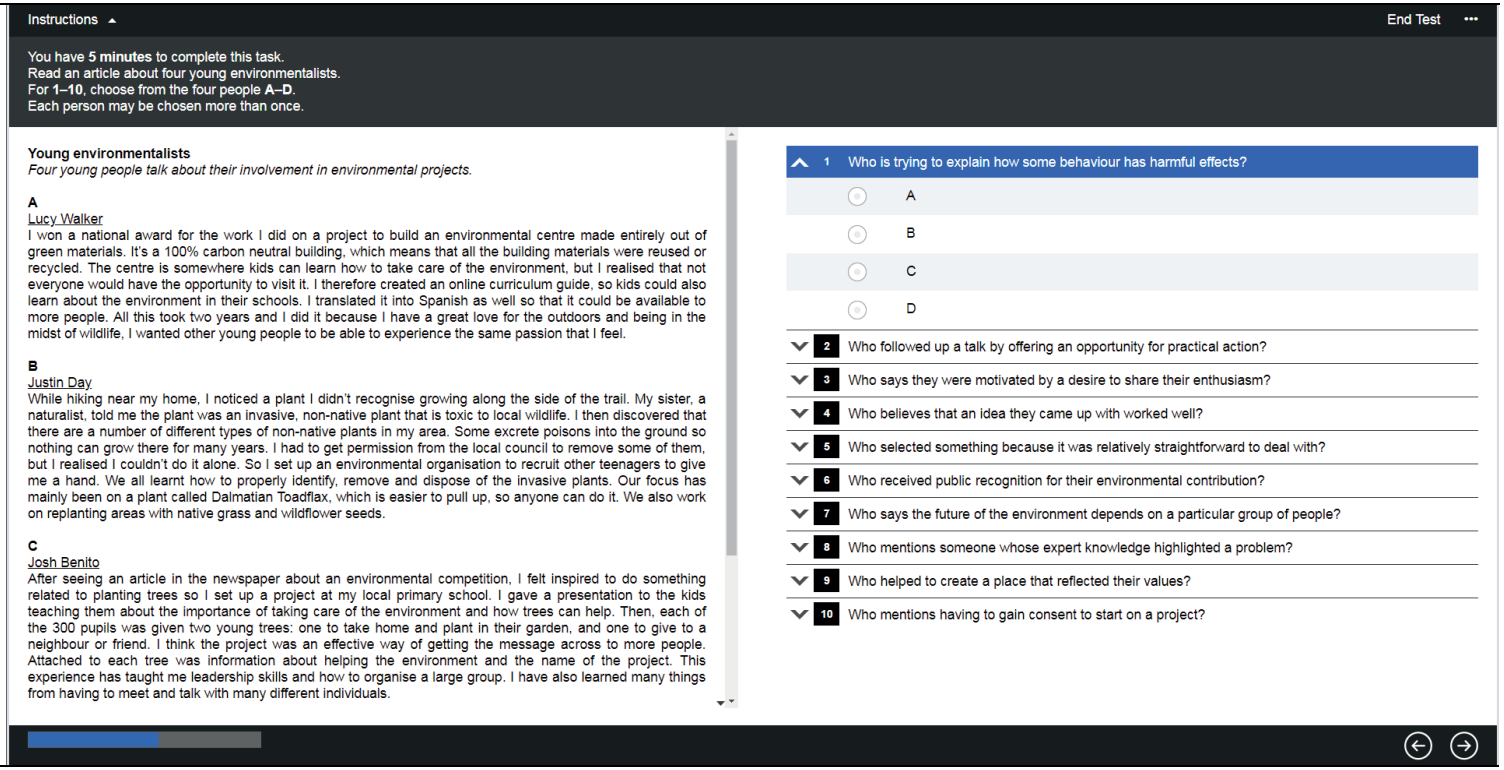
Task type	CEFR level	Sample task
Matching sentences	**A1**	

Task type	CEFR level	Sample task
Discrete multiple-choice	A1 **A2** B1	

Task type	CEFR level	Sample task
Multi-item, multiple-choice	**B1** B2 C1	

Task type	CEFR level	Sample task
Matching gapped text - sentences	**B1**	

Task type	CEFR level	Sample task
Matching gapped text - paragraphs	**C1**	

Task type	CEFR level	Sample task
Cross-text matching	**B2**	

Source: Cambridge Assessment English.

Annex 4.C. Mapping of listening tasks to framework dimensions

Annex Table 4.C.1 maps the PISA Foreign Language Assessment listening tasks to the relevant Common European Framework of Reference for Languages (CEFR) can-do statements. In addition to the CEFR mapping, the table captures task- and CEFR level information in relation to the cognitive processes involved and task description. Points to note are as follows:

- The *Overall oral comprehension* descriptor is included for each task and level to act as a reminder of the broad listening ability in action at each level. Certain other scales and descriptors frequently appear across the table, e.g. the scale *Identifying cues and inferring.* These frequently-represented scales, and the selected descriptors from these scales, highlight those abilities employed widely across a range of listening task types. An example of a descriptor from the above-mentioned scale at Level A1 is *Can deduce the meaning of an unknown word for a concrete action or object, provided the surrounding [spoken] text is very simple, and on a familiar everyday subject.* This can-do statement will have general applicability to a number of different task types and, as such, appears in the table in several places.
- Some scales feature less frequently across the table due to their more specific relevance to certain task types, e.g. *Understanding announcements and instructions.*
- The featured descriptors from the relevant scales highlight the particular abilities used in specific tasks at certain levels for the successful completion of that listening task.
- Some scales are not represented on the table of listening tasks (e.g. *Watching TV, film and video*) as these are not currently part of the test construct.
- Descriptors from plus-levels are not included in the table.
- The highest level of cognitive processing required for each task is included, with the assumption that the other lower-level process will also be utilised.
- Please note that aspects of the can-do statements that are not relevant to the PISA cohort have been struck through.

Annex Table 4.C.1. Mapping of listening tasks to CEFR level, cognitive process and CEFR can-do statements

Task type	CEFR level	Task description	Cognitive process	CEFR can-do statements
Listening – Discrete multiple-choice with graphics	Pre-A1	**Audio**: Short informal or neutral dialogue **Response format**: Question and three picture options **Operation**: Select the correct picture according to the audio	Input decoding	**Overall oral comprehension** Can understand short, very simple questions and statements, provided they are delivered slowly and clearly and accompanied by visuals or manual gestures to support understanding, and repeated if necessary. Can recognise everyday, familiar words~~/signs~~, provided they are delivered clearly and slowly in a clearly defined, familiar everyday context. Can recognise numbers, prices, dates and days of the week, provided they are delivered slowly and clearly in a defined, familiar everyday context. **Identifying cues and inferring** Can deduce the meaning of a word~~/sign~~ from an accompanying picture or icon.
	A1	**Audio**: Short informal or neutral dialogue **Response format**: Question and three picture options **Operation**: Select the correct picture according to the audio	Lexical search	**Overall oral comprehension** Can follow language that is very slow and carefully articulated, with long pauses for them to assimilate meaning. Can recognise concrete information (e.g. places and times) on familiar topics encountered in everyday life, provided it is delivered slowly and clearly. **Understanding conversation between other people** Can understand some expressions when people are discussing them, family, school, hobbies or surroundings, provided the delivery is slow and clear. Can understand words~~/signs~~ and short sentences in a simple conversation (e.g. between a customer and a salesperson in a shop), provided people communicate very slowly and very clearly. **Understanding announcements and instructions** Can understand instructions addressed carefully and slowly to them and follow short, simple directions. Can understand when someone tells them slowly and clearly where something is, provided the object is in the immediate environment. **Identifying cues and inferring** Can deduce the meaning of an unknown word~~/sign~~ for a concrete action or object, provided the surrounding [spoken] text is very simple, and on a familiar everyday subject.
	A2	**Audio**: Short informal or neutral dialogue **Response format**: Question and three picture options **Operation**: Select the correct picture according to the audio	Parsing	**Overall oral comprehension** Can understand phrases and expressions related to areas of most immediate priority (e.g. very basic personal and family information, shopping, local geography~~, employment~~), provided people articulate clearly and slowly. **Understanding conversation between other people** Can follow in outline short, simple social exchanges conducted very slowly and clearly. **Understanding announcements and instructions**

Task type	CEFR level	Task description	Cognitive process	CEFR can-do statements
				Can understand simple directions on how to get from X to Y, by foot or public transport. Can understand basic instructions on times, dates and numbers, etc., and on routine tasks and assignments to be carried out. **Identifying cues and inferring** Can exploit numbers, dates, names, proper nouns, etc. to identify the topic of a text.
Listening – Discrete multiple-choice with text	B1	**Audio**: Range of input – dialogue, monologue, news report, interview **Response format**: Question/statement and three text-based options **Operation**: Select the correct option according to the audio	Parsing	**Overall oral comprehension** Can understand the main points made in clear standard language or a familiar variety on familiar matters regularly encountered at ~~work,~~ school, leisure, etc., including short narratives. **Understanding conversation between other people** Can generally follow the main points of extended discussion around them, provided it is clearly articulated in standard language or a familiar variety. **Understanding audio media and recordings** Can understand the main points of news bulletins and simpler recorded material about familiar subjects delivered relatively slowly and clearly. Can understand the main points and important details in stories and other narratives (e.g. a description of a holiday), provided the delivery is slow and clear. **Identifying cues and inferring** Can follow a line of argumentation or the sequence of events in a story, by focusing on common logical connectors (e.g. however, because) and temporal connectors (e.g. after that, beforehand).
	B2	**Audio**: Range of input – dialogue, monologue, news report, interview **Response format**: Question/statement and three text-based options **Operation**: Select the correct option according to the audio	Meaning construction	**Overall oral comprehension** Can follow extended discourse and complex lines of argument, provided the topic is reasonably familiar, and the direction of the argument is signposted by explicit markers. **Understanding conversation between other people** Can identify the main reasons for and against an argument or idea in a discussion conducted in clear standard language or a familiar variety. Can follow a chronological sequence in extended informal discourse, e.g. in a story or anecdote. **Understanding audio media and recordings** Can understand most documentaries and most other recorded or broadcast material delivered in the standard form of the language and can identify mood, attitude, etc. **Identifying cues and inferring** Can use a variety of strategies to achieve comprehension, including watching out for main points and checking comprehension by using contextual clues.
	C1	**Audio**: Range of input – dialogue, monologue, news report, interview **Response format**: Question/statement and three text-based options	Meaning construction	**Overall oral comprehension** Can understand enough to follow extended discourse on abstract and complex topics beyond their own field, though they may need to confirm occasional details, especially if the variety is unfamiliar. Can recognise a wide range of idiomatic expressions and colloquialisms, appreciating register shifts.

Task type	CEFR level	Task description	Cognitive process	CEFR can-do statements
		Operation: Select the correct option according to the audio		Can follow extended discourse even when it is not clearly structured and when relationships are only implied and not signalled explicitly. **Understanding conversation between other people** Can easily follow complex interactions between third parties in group discussion and debate, even on abstract, complex, unfamiliar topics. Can identify the attitude of each participant in an animated discussion characterised by overlapping turns, digressions and colloquialisms that is delivered at a natural speed in varieties that are familiar. **Understanding audio media and recordings** Can understand a wide range of recorded and broadcast material, including some non-standard usage, and identify finer points of detail including implicit attitudes and relationships between people. **Identifying cues and inferring** Is skilled at using contextual, grammatical and lexical cues to infer attitude, mood and intentions and anticipate what will come next.
Listening – Multi-item multiple-choice	B1	**Audio**: Range of input – dialogue, extract from a news programme, interview **Response format**: Five questions / statements and three text-based options **Operation**: Select the correct option according to the audio	Parsing	**Overall oral comprehension** Can understand the main points made in clear standard language or a familiar variety on familiar matters regularly encountered at ~~work,~~ school, leisure, etc., including short narratives. **Understanding conversation between other people** Can generally follow the main points of extended discussion around them, provided it is clearly articulated in standard language or a familiar variety. **Understanding audio media and recordings** Can understand the main points of news bulletins and simpler recorded material about familiar subjects delivered relatively slowly and clearly. Can understand the main points and important details in stories and other narratives (e.g. a description of a holiday), provided the delivery is slow and clear. **Identifying cues and inferring** Can follow a line of argumentation or the sequence of events in a story, by focusing on common logical connectors (e.g. however, because) and temporal connectors (e.g. after that, beforehand).
	C1	**Audio**: Range of input – dialogue, extract from a news programme, interview **Response format**: 1) Five questions / statements and four text-based options 2) Two questions/statements and three text-based options **Operation**: Select the correct option according to the audio	Meaning construction Discourse construction	**Overall oral comprehension** Can understand enough to follow extended discourse on abstract and complex topics beyond their own field, though they may need to confirm occasional details, especially if the variety is unfamiliar. Can recognise a wide range of idiomatic expressions and colloquialisms, appreciating register shifts. Can follow extended discourse even when it is not clearly structured and when relationships are only implied and not signalled explicitly. **Understanding conversation between other people**

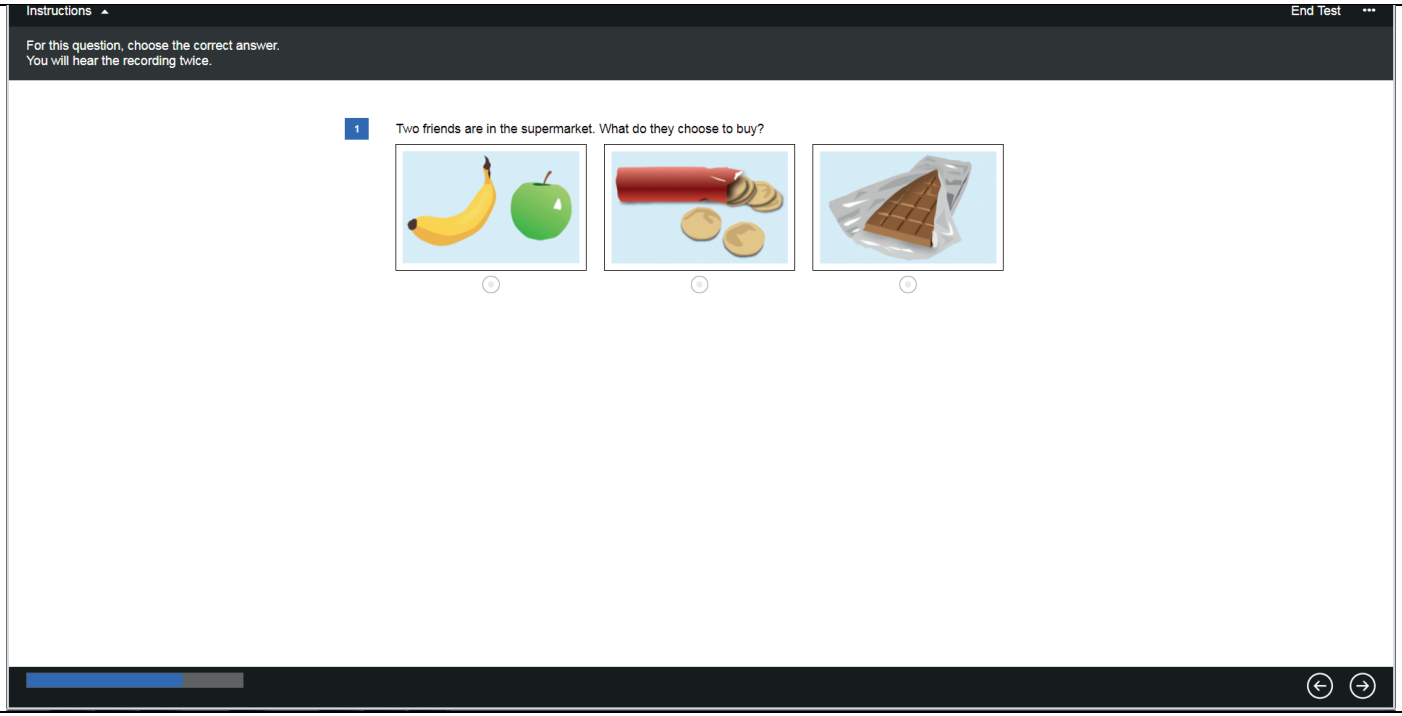
Task type	CEFR level	Task description	Cognitive process	CEFR can-do statements
		Input text: Long text		Can easily follow complex interactions between third parties in group discussion and debate, even on abstract, complex, unfamiliar topics. Can identify the attitude of each participant in an animated discussion characterised by overlapping turns, digressions and colloquialisms that is delivered at a natural speed in varieties that are familiar. **Understanding audio media and recordings** Can understand a wide range of recorded and broadcast material, including some non-standard usage, and identify finer points of detail including implicit attitudes and relationships between people. **Identifying cues and inferring** Is skilled at using contextual, grammatical and lexical cues to infer attitude, mood and intentions, and anticipate what will come next.

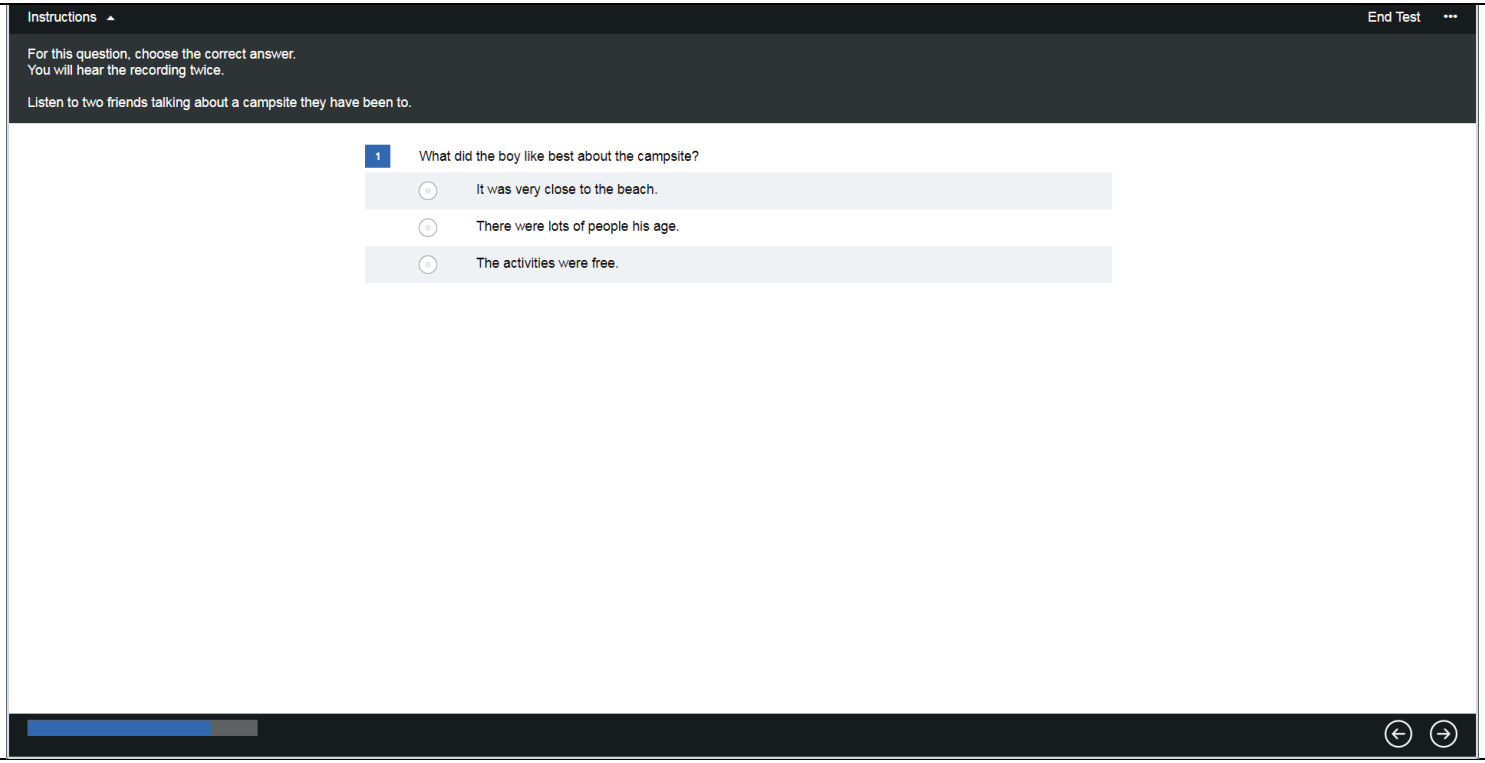
Source: Cambridge Assessment English's original work, based on Council of Europe (2020[32]), *Common European Framework of Reference for Languages: Learning, teaching, assessment – Companion volume*, © Council of Europe, Strasbourg, pp. 48-52 and Field, J. (2013[40]), *Cognitive validity*, in Taylor, L. and A. Geranpayeh (eds.), *Examining Listening: Research and Practice in Assessing Second Language Listening*, UCLES/Cambridge University Press, Cambridge, pp. 95-103.

Annex 4.D. Sample tasks: Listening

An example is given for each task type. Please note the same task type may be used at a number of different levels. Where a task type is used at more than one level, the level of the example is shown in bold.

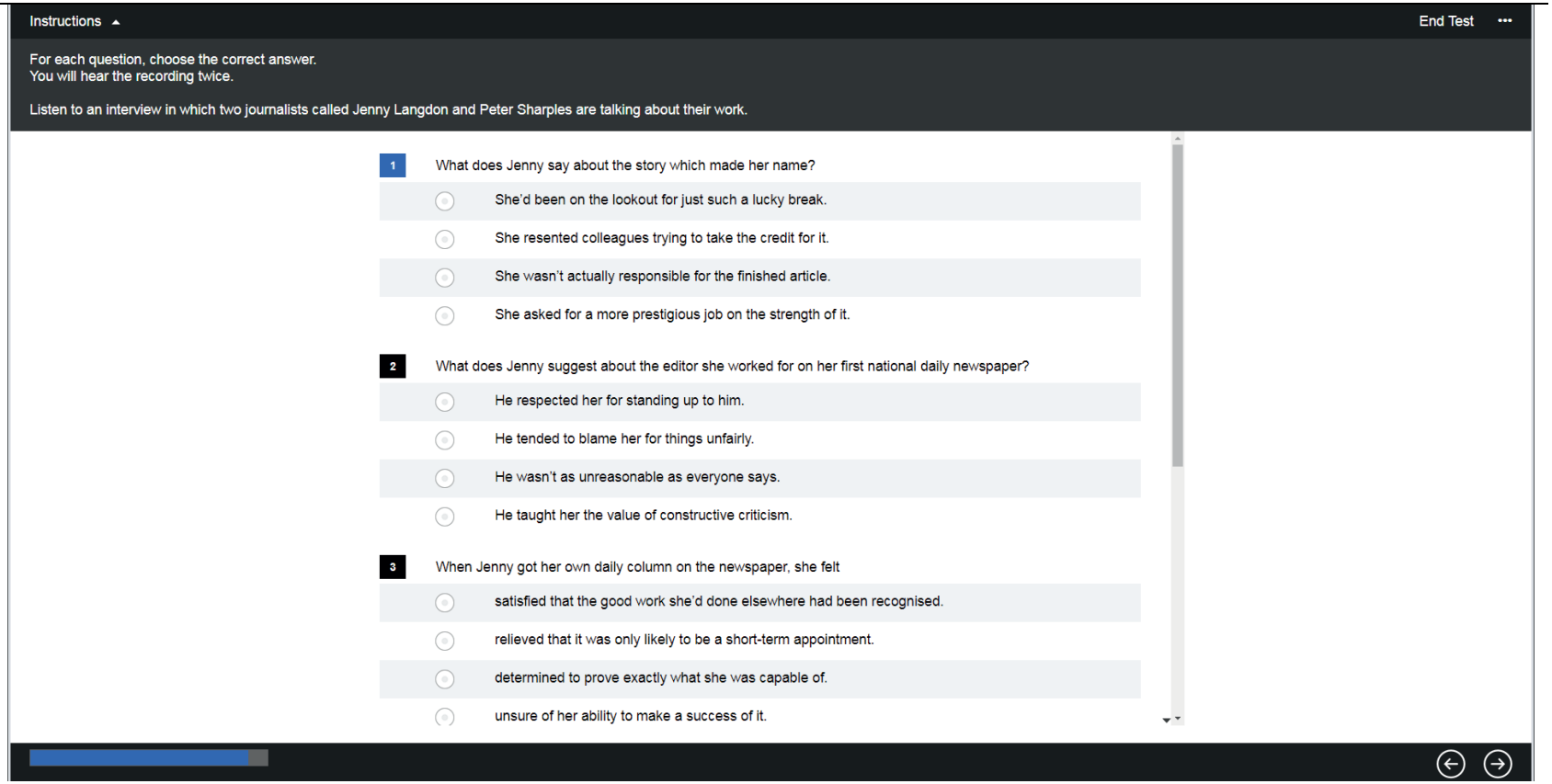
Annex Table 4.D.1. Sample listening tasks

Task type	CEFR level	Sample task	
Discrete multiple choice with graphics	Pre-A1 **A1** A2		**SCRIPT** **M=Male;** **F=Female** **M**: Let's buy something to eat. **F**: OK. What do you want? Some chocolate? **M**: Sorry, I don't like it. What about fruit? **F**: Yes, let's get some. And I've got some cookies in my bag.

Task type	CEFR level	Sample task
Discrete multiple-choice with text	**B1** B2 C1	

SCRIPT
M=Male; F=Female

M: So, how did you like the campsite?
F: Brilliant! Thanks for telling my dad about it! It's got everything — shops, swimming pool, games room. And a beautiful beach five minutes away.
M: I knew you'd love that. The pool's good enough for me. I had the best time ever — so much to do!
F: When we first got there, I felt a bit lost. But then I met some kids in the same school year as me.
M: That's why I'd go back any day — in other camps you don't get that. I remember one where it was mostly families with little kids.
F: I know what you mean.

Task type	CEFR level	Sample task
Multi-item multiple-choice	B1 **C1**	

SCRIPT
Int=Interviewer

Int: Today we're looking at careers in journalism. My guests are Jenny Langdon and Peter Sharples, both regular columnists on major publications. Jenny, you made your name really young, didn't you?

F: Relatively, yes. I was a raw recruit on the local paper when I uncovered a story concerning a celebrity living nearby. Out of the blue I found myself with a scoop on my hands. Basically, I found the guy,

Task type	CEFR level	Sample task
		interviewed him, then hid him some place where reporters on rival papers wouldn't find him. When the story broke next day, the editorial team had actually cobbled the front-page story together from my notes, but it was attributed to me by name. Before I knew what was happening, I'd been headhunted by a national daily. It was a turning point alright – but I can hardly claim it as a shrewd career move or anything! **Int**: And the editor at that national daily was a notoriously bad-tempered individual … **F**: Well, there's no denying he deserved that reputation! I mean, having landed a dream job, I was really thrown in at the deep end! My desk was right outside his office, so I was first in the firing line if anything went wrong – even stuff I'd had no hand in! But I knew better than to argue, and was thick-skinned enough not to take it personally. Anyway that's what the paper was like, always on the edge, and I really flourished in that environment. **Int**: Eventually getting your own daily column … **F**: … and that's where I really came into my own. I mean, I'd done stints on the sports desk, been celebrity correspondent – the works. Actually, I only got offered the column as a stop-gap when my predecessor left under a cloud. But I was desperate to hold on to it. And it came at just the right time – if it'd been earlier, I'd never have had the nerve or the experience to make it my own. **Int**: Let's bring Peter in here. You started off on the celebrity magazine called 2U, didn't you? **M**: I did. Ostensibly thanks to a speculative letter to the editor when I was still a student. Actually, I'd been doing stuff for a student newspaper all through university. Skills I learnt there stood me in good stead. When 2U Magazine called me for interview, my approach to college news convinced them I was in touch with reality – you know, budgets, deadlines, all that – that's what swung it in my favour – it wasn't just having my finger on the pulse as far as youth culture was concerned – important as that was at 2U. **Int**: Can I ask you both whether you'd say courses in journalism are worth doing? Jenny? **F**: Well, I wanted to write and a journalism course seemed a reasonable enough starting point. Journalism is at least paid up front – unlike some forms of writing, and there's no denying that was an incentive. So, yes, I did one. And, you know, if I hadn't, who knows if I'd have been able to handle the stuff thrown at me when I first arrived at the newspaper – it does give you that grounding. But I wouldn't say it taught me everything I needed. Fortunately a stint on the student newspaper filled in the gaps. **M**: … as is so often the case. They're often criticised for taking too strong a line on issues, but they're invaluable because they give you that free rein, and you're generally writing from the heart rather than for the money. I'd say by all means do a course, theorise all you like in the classroom, but just bear in mind that it's no substitute for getting out there – for developing your own style.

Source: Cambridge Assessment English.

Annex 4.E. Mapping of speaking tasks to framework dimensions

Annex Table 4.E.1 and Source: Cambridge Assessment English's original work, based on Council of Europe , *Common European Framework of Reference for Languages: Learning, teaching, assessment- Companion volume*, © Council of Europe, Strasbourg, p.62.

Annex Table 4.E.2 map PISA Foreign Language Assessment speaking tasks to the relevant overall oral production and overall phonological Common European Framework of Reference (CEFR) can-do statements. Points to note are as follows:

- Due to the controlled natured of the output, Part 1 (Read aloud) is only mapped to the overall phonological CEFR scale.

Annex Table 4.E.1. Mapping speaking tasks to CEFR level and overall oral production can-do statements

Task type	CEFR level	Overall oral production can-do statements
Part 2 – Odd-one out Part 3 – Long turn – picture Part 4 – Storytelling Part 5 – Long turn	Pre-A1	Can produce short phrases about themselves, giving basic personal information (e.g. name, address, family, nationality).
	A1	Can produce simple, mainly isolated phrases about people and places.
	A2	Can give a simple description or presentation of people, living or working conditions, daily routines, likes and dislikes, etc. as a short series of simple phrases and sentences linked into a list.
	B1	Can reasonably fluently sustain a straightforward description of one of a variety of subjects within their field of interest, presenting it as a linear sequence of points.
	B2	Can give clear, detailed descriptions and presentations on a wide range of subjects related to their field of interest, expanding and supporting ideas with subsidiary points and relevant examples.
	C1	Can give clear, detailed descriptions and presentations on complex subjects, integrating subthemes, developing particular points and rounding off with an appropriate conclusion.

Source: Cambridge Assessment English's original work, based on Council of Europe (2020[32]), *Common European Framework of Reference for Languages: Learning, teaching, assessment- Companion volume*, © Council of Europe, Strasbourg, p.62.

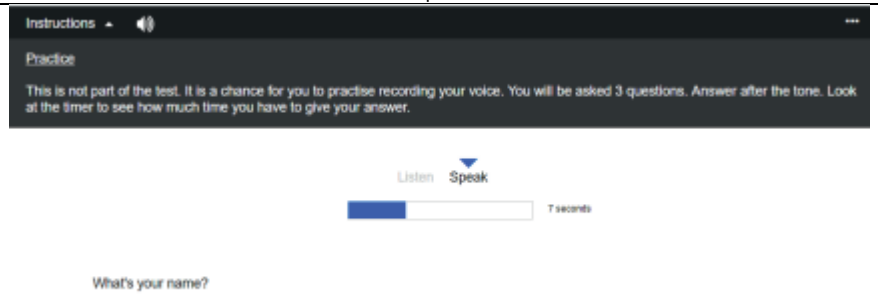
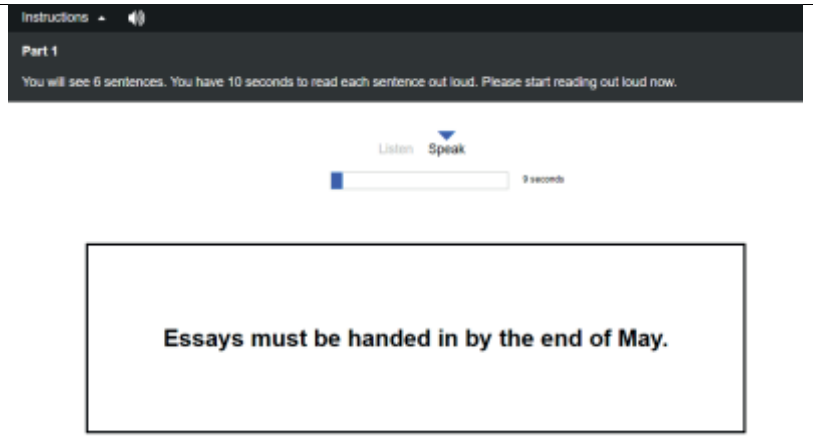
Annex Table 4.E.2. Mapping speaking tasks to CEFR level and overall phonological control can-do statements

Task type	CEFR level	Overall phonological control can-do statements
Part 1 – Read aloud Part 2 – Odd-one out Part 3 – Long turn – picture Part 4 – Storytelling Part 5 – Long turn	A1	Pronunciation of a very limited repertoire of learned words and phrases can be understood with some effort by interlocutors used to dealing with speakers of the language group. Can reproduce correctly a limited range of sounds as well as stress for simple, familiar words and phrases.
	A2	Pronunciation is generally clear enough to be understood, but conversational partners will need to ask for repetition from time to time. A strong influence from the other language(s) they speak on stress, rhythm and intonation may affect intelligibility, requiring collaboration from interlocutors. Nevertheless, pronunciation of familiar words is clear.
	B1	Pronunciation is generally intelligible; intonation and stress at both utterance and word levels do not prevent understanding of the message. Accent is usually influenced by the other language(s) they speak.
	B2	Can generally use appropriate intonation, place stress correctly and articulate individual sounds clearly; accent tends to be influenced by the other language(s) they speak, but has little or no effect on intelligibility.
	C1	Can employ the full range of phonological features in the target language with sufficient control to ensure intelligibility throughout. Can articulate virtually all the sounds of the target language; some features of accent(s) retained from other language(s) may be noticeable, but they do not affect intelligibility.

Source: Cambridge Assessment English's original work, based on Council of Europe (2020[32]), *Common European Framework of Reference for Languages: Learning, teaching, assessment- Companion volume*, © Council of Europe, Strasbourg, pp. 134-135.

Annex 4.F. Sample tasks: Speaking

Annex Table 4.F.1. Sample speaking tasks

Task type	Sample task
Practice	
Part 1 - Read aloud	

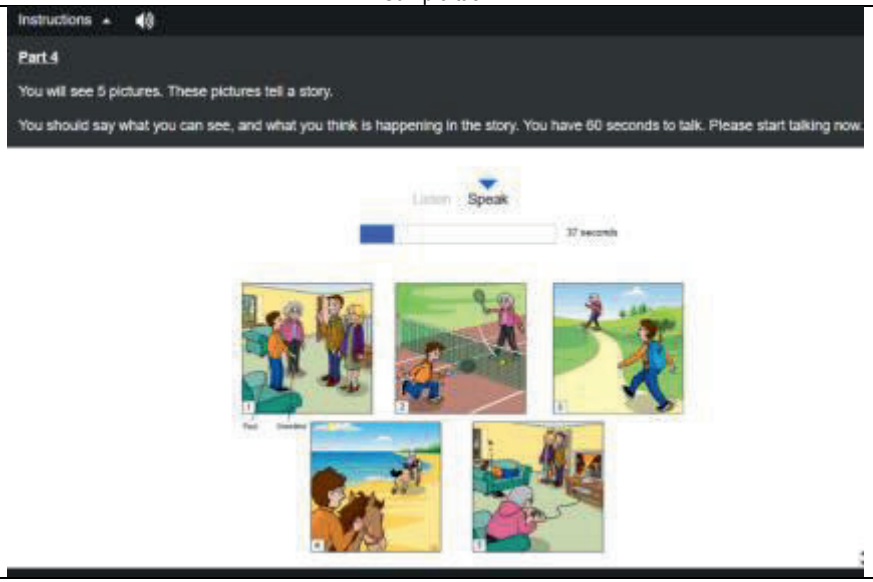
Task type	Sample task
Part 2 – Odd- one out	**Part 2** You will see 4 pictures. One of them is different. First you have 5 seconds to look at the pictures. Please say which picture is different, and why it is different. There are 4 sets of pictures. 
Part 3 - Long turn - picture	

Task type	Sample task

Part 3

You will have to talk about a picture. First, you have 10 seconds to look at the picture. You will be asked 2 questions. You have 30 seconds for each question.

Question 1: What is happening in the picture?
Question 2: Why do you think the people are smiling?

Task type	Sample task
Part 4 - Storytelling	

Task type	Sample task
Part 5 - Long turn	**Part 5** I'd like you to talk about learning English What are the best ways to learn English? Here are some ideas to help you. You now have 1 minute to talk about what are the best ways to learn English and why?. 

Source: Cambridge Assessment English.

5 Background questionnaires framework

The PISA foreign language assessment will include a set of context questionnaires. This chapter sets the overall directions for questionnaire development, based on an in-depth review of the scientific literature, a review of past large-scale international assessments, and input from experts and policy makers. The chapter provides a comprehensive picture of the factors influencing foreign language learning and proficiency inside and outside school, at the student, parent, teacher, school and system levels. The policy and contextual information is structured around a set of constructs that are grouped in four policy domains: government and school policies; teacher training and profiles; students and learning; and teaching practices.

Contents and methodology

Overview

For the first time, the PISA 2025 cycle will provide policy makers and educators with an assessment of 15-year-old students' foreign language proficiency.

An important goal of the PISA foreign language assessment (FLA) is to compare how students around the world learn foreign languages and identify best practices in the teaching and learning of foreign languages in school. This will contribute to the improvement of foreign language teaching and learning, and guide policy decisions.

To accomplish this goal, the PISA FLA will include a set of questions for the PISA teacher, student, school and parent questionnaires, and a system-level questionnaire asking for information on foreign language teaching and learning from policy officials.

This chapter presents the framework for the PISA 2025 FLA background questionnaires. It illustrates the policy and contextual information that can be used to interpret the data on students' proficiency in foreign languages and produce policy-relevant analyses, structured around a set of constructs. Based on this framework, the questionnaires were developed (between January and December 2020) following a similar process of consultation with experts and country delegates. A list of the contributing experts and their institutional affiliations can be found in Annex 5.C.

This framework determined the overall directions for questionnaire development. However, due to space limitations not all constructs in this chapter have been included in the PISA background questionnaires. To help guide decisions on the coverage of each construct in the questionnaires, every construct in the framework is rated as relevant or highly relevant for policy, based on country feedback.

Box 5.1. Terminological notes

Throughout this chapter, the following terms will be used:

- foreign languages are all modern languages that are formally taught in school settings, other than the main language of schooling, which is defined as the language of the PISA reading test.
- the target language is the language of the PISA foreign language assessment; in PISA 2025, the target language will be English, but using a general terminology will make it easier to adapt this chapter to other languages that may be tested in the future.
- the reading test language is the language of the PISA reading assessment.
- other foreign languages are other languages that students study at school and are different from the target and reading test languages.

The skills of reading comprehension, spoken production, listening comprehension and written production will be referred to as reading, speaking, listening and writing, respectively.

Following the Common European Chapter of Reference for Languages (CEFR) (Council of Europe, 2001, pp. 4-5[1]), multilingualism is defined in this chapter as "the knowledge of a number of languages, or the co-existence of different languages in a given society". Plurilingualism is defined as the ability to communicate effectively with a particular interlocutor, simultaneously using a variety of linguistic and cultural skills to do so.

Chapter content

The context questionnaires chapter is divided into four main policy domains Figure 5.1. Each policy domain contains the sections presented in the second part of this chapter, which, in turn, include a number of constructs. The categorisation of the constructs into four domains (which is to some extent arbitrary) is only meant to simplify the exposition; it does not affect the description and policy relevance of the constructs or the development of the questionnaire. For example, "languages learned at school" is included in the setting for target language teaching in Figure 5.1 because it includes languages as compulsory subjects and also languages used as a medium to learn other subjects. However, the choice of which languages to study is primarily a student choice in many education systems. Therefore, the policy domains should not be seen as mutually exclusive, but rather as interconnected (for example, through the constructs they potentially share).

Each domain can include both policy levers over which schools and governments have direct control and contextual/external factors defining or constraining policy over which schools and governments have little control (OECD, 2003[2]; OECD, 2018, p. 14[3]). The main outcome of interest is foreign language proficiency, and the main goal of the analysis will be to relate proficiency to policy levers, controlling for the relevant contextual factors. However, other constructs in this chapter can also be (or contain) outcomes of interest in themselves (e.g. respect and openness towards people from other culture and language background, construct (22); and intrinsic motivation to learn the target language, see construct (17)).

Figure 5.1. The context questionnaires chapter

Policy domain 1: Government and school policies
The setting for target language learning at school
(1) Onset of target language learning at school
(2) Intensity of target language learning at school
(3) Target language class size
(4) Languages learned at school
The school environment
(5) Availability of foreign languages
(6) School enrichments activities for target language learning
(7) Target language remedial lessons at school
(8) School resources for foreign language teaching

Policy domain 3: Teachers' training and profile
Human resources
(24) Target language teaching experience
(25) Target language teacher and staff availability
(26) Teachers' initial education and qualifications
(27) Teachers' in-service training
(28) Teaching specialisation
The teacher
(29) Teachers' visits to other language communities
(30) Teachers' attitudes related to target language teaching
(31) Teachers' target language proficiency

Policy domain 2: Students and learning
Student background, environment and family support
(9) Language background
(10) Parents' target language proficiency
(11) Family support in target language learning
(12) Family and peers' perceptions and attitudes related to the target language
(13) Family and peers' perceptions and attitudes related to target language lessons
(14) Target language exposure through the media
(15) Students' visits to other language communities
(16) Face-to-face exposure to and use of target and foreign languages outside of school
Learner's attitudes, motivations and behaviours
(17) Students' motivation for target language learning
(18) Students' perceived level of proficiency
(19) Students' attitudes towards target language learning at school
(20) Time spent on target language study, homework and other structured learning activities outside of school
Intercultural and multilingual environments and target language learning
(21) Linguistic and cultural diversity in the community
(22) Respect and openness towards people from other culture and language backgrounds
(23) Pluricultural and plurilingual education

Transversal topics
Information and Communication Technologies
Use of the target language for instruction in other subjects

Policy domain 4: Teaching practices
Teaching approaches and methods
(32) Teaching the four communicative skills
(33) Teaching linguistic knowledge: Grammar, pronunciation, vocabulary
(34) Teaching literature and cultural knowledge
(35) Group and collective learning activities in the classroom
(36) Translanguaging
(37) Use of the target language during foreign language lessons
(38) Teacher's talking time
(39) Joint learning of language- and non-language-related content
(40) Teaching materials used for target language teaching
(41) Use of teaching materials
(42) Use of the Common European Framework of Reference for Languages
Assessment practices
(43) Existence of system-level target language assessments
(44) Assessment for learning

Source: Authors

The policy domain “government and school policies” presents a broad characterisation of target language learning in an education system and its schools. It encompasses the general setting for foreign language learning at school (e.g. target language teaching onset and teaching time) and the school environment (e.g. school resources and activities). The constructs in this domain can be strongly influenced by the government regulatory chapter and by school decisions, although they also depend on the interaction between these regulations and other contextual and policy factors.

“Students and learning” reviews a variety of student characteristics and behaviours, and their association with other constructs in the chapter and with foreign language proficiency and learning. It contains constructs related to the students’ background and the environment where they live (e.g. language background, family support and daily exposure to the target language outside of school); their motivations and attitudes (e.g. motivation to learn the target language and attitudes towards other cultures), and their behaviours (e.g. time spent on homework or engagement with a variety of media in the target language). Many of the constructs included in this section represent contextual factors that are difficult to modify (e.g. language background), but there are also areas more susceptible to policy intervention (e.g. remedial lessons).

“Teachers’ training and profile” relates to the characteristics of the target language teaching workforce and to the main policies to train and employ this workforce. For example, the section on human resources addresses issues of training, qualifications and the specialisation of teachers. The section on the teacher contains constructs such as “teachers’ target language proficiency” and “teachers’ visits to other language communities”. Governments usually have some control over this policy domain, either in the short term (e.g. in-service training) or in the longer term (e.g. criteria to become a target language teacher, affecting staff availability in the course of years).

“Teaching practices” refers to what happens in the classroom. Governments and schools can influence this domain through guidelines and recommendations. However, their ability to influence teaching practices may depend on the implementation of their policies and regulations. Teaching practices include the use of broad teaching approaches (e.g. communicative language teaching) as well as more specific methods (e.g. letting students work in groups) and evaluation practices.

The remainder of this section presents in more detail the scope of this context questionnaires chapter and the methodology used to develop it. The second section of this chapter presents the constructs depicted in Figure 5.1.

Relationship with the main PISA questionnaire framework

Questions related to foreign language learning and proficiency will be asked in addition to the other questions included in the PISA context questionnaires. The current PISA context questionnaires framework (OECD, 2019[4]) indicates a set of constructs, divided into domain-general (also called more briefly “general”) and domain-specific. General constructs are important for understanding differences in achievement that are not tied to a specific subject area. Domain-specific constructs are those with a strong expected relationship to student experiences, outcomes, and teaching and learning factors tied to a specific content area.

The foreign language context questionnaires chapter covers domain-specific constructs tied to foreign language learning. General constructs (possibly affecting foreign language learning) are discussed in the PISA context questionnaires chapter (OECD, 2019[4]) and (when referring to teachers) in the TALIS (the OECD Teaching and Learning International Survey) conceptual chapter (Ainley and Carstens, 2018[5]). Therefore, general constructs (e.g. student socio-economic status) are not covered in this chapter, except when a specific and direct relationship with foreign language learning (e.g. language background) is suggested.

Methodology

The constructs included in this chapter were identified and defined through a four-step process. First, the theoretical chapter and/or the questionnaires for the main international comparative studies of foreign language learning that were conducted (or planned) in the past have been studied to identify relevant domain-specific concepts (Table 5.1). Of these studies, SurveyLang (European Commission, 2012[6]) was of particular significance as it was conducted recently and its proposed instruments, target age and competencies assessed are similar to those of the PISA foreign language assessment.

Table 5.1. International comparative studies of foreign language learning

Survey	Data collection years	Institution	Participating systems	Age group	Languages assessed	Type of instruments
The survey of English as a foreign language (The IEA Six-subject survey) (Lewis and Massad, 1975[7])	1970-1971	International Association for the Evaluation of Educational Achievement (IEA)	10	14-year-old students and students in the final year of lower secondary school	English*	Proficiency test (reading, listening, writing, speaking) Questionnaires (students, teachers, school principals)
The Language Education Study (Peter Dickson and Alister Cumming, 1996[8]; IEA, 1993[9])	1990 (partial execution)	IEA	25	15-18 year-old students (planned)	English, French, German and Spanish	Proficiency test (never conducted) Questionnaire (system-level)
The European Survey on Language Competences (SurveyLang) (European Commission, 2012[6])	2011	European Commission	16	Lower (final year) or upper (second year) secondary students	English, French, German, Italian and Spanish	Proficiency test (reading, listening, writing) Questionnaires (students; teachers; school principals; system-level)
The Early Language Learning in Europe (ELLiE) (Enever, 2011[10])	2007-2010	British Council	7	7-8 year-old students in primary education	English, French and Spanish	Qualitative interviews (students; teachers; school principals) Questionnaire (students; parents; system-level)
The Eurydice Key Data on Teaching Languages at School in Europe (European Commission/EACEA/Eurydice, 2017[11])	2017	European Commission	34	Primary to upper secondary education (ISCED1 to ISCED3)	All languages	Questionnaire (system-level)

Note: The IEA Six-subject survey also included an assessment of French (Carroll, 1975[12]) with eight participating education systems. The set-ups of the English and French assessment studies present a substantial degree of similarity (Cumming, 1996[13]), due to the close cooperation between the committees developing them, and the decision to adopt the same basic design (Lewis and Massad, 1975[7]).

Second, a review of the scientific literature was conducted to refine the constructs and to check whether some important constructs had been overlooked by the projects listed above. A first version of the context questionnaires chapter was completed describing the constructs identified through the first two steps.

Third, the constructs were revised based on the feedback received from national representatives from 13 PISA countries and from 5 independent experts with substantial expertise in the learning of foreign languages. These experts gave written feedback on the chapter, and subsequently met in Paris to discuss the chapter and feedback on 15-17 May 2019.

Fourth, the context questionnaires chapter was further revised based on the feedback received from an extended group of 14 independent experts with substantial expertise in the learning of foreign languages; an internal peer review within the OECD; and further inputs from country delegates participating in the PISA Governing Board meeting in September 2019.

Determination of the policy relevance of the constructs

As seen in the previous sections, this chapter covers a wide range of topics. However, the questionnaires will have considerable length restrictions. Therefore, the policy relevance of the constructs will be one of the deciding factors when choosing which constructs to incorporate into the questionnaires, together with technical criteria related to questionnaire design. PISA-participating countries and economies will be involved in selecting the constructs to incorporate into the questionnaires.

This chapter proposes a tentative classification of the policy relevance of the constructs as "essential" or "relevant" (constructs deemed not relevant for policy analysis are excluded).[1] The policy relevance was derived in the following way:

- the authors identified and described in the first draft of this chapter a list of constructs affecting foreign language proficiency and learning.
- a survey was sent to all PISA-participating countries/economies to rate the relevance and comment on the constructs included in the first draft of this chapter; five independent experts gave additional comments.
- experts and representatives of countries met in Paris and, based on the aggregate country ratings, assessed the policy relevance of each construct; the resulting policy relevance is reported in this chapter.[2]

The decision on the policy relevance of each construct was reached at the meeting by consensus, meaning that all participants agreed with the decision. In total, 13 countries rated the constructs or participated to the assessment of the constructs' relevance at the meeting in Paris.[3]

Structure of the chapter presentation

The following sections of this chapter discuss the policy domains, each containing a number of subsections associated to a policy question and a group of constructs as illustrated in Figure 5.1. For each construct discussed in the following sections, the following elements are reported:

- a basic description
- the way the construct can be expected to be associated with foreign language learning or to other constructs in the chapter
- the levels at which the construct can be measured (student, parent, teacher, school or system)
- a policy relevance rating based on the preferences expressed by PISA-participating countries/economies interested in the PISA foreign language assessment

The latter two elements (measurement level and policy relevance) are reported in a summary table below the introduction of each section.

The suggested measurement levels are based on the description of each construct and on the opinion of the experts who reviewed the chapter. A construct could be measured at one level, but also (when allowed by space constraints) at multiple levels. The latter option could allow for an assessment of the correspondence between policies and practices, or for improving the accuracy of the measurement by using multiple measures to capture the same construct ("triangulation" (Heath, 2015, p. 639[14])).

This chapter does not offer specific recommendations on this issue, and the suggested measurement levels are only indicative.

Transversal topics: ICT and the use of the target language for instruction in other subjects

Language learning across countries and economies is affected by broad trends that interact with the constructs included in this chapter and contribute to shape them. Two particularly important ones are the diffusion of ICT and the use of foreign languages for instruction in other subjects. These two broad trends have been considered transversal topics, and integrated in the discussion of constructs throughout this chapter.

Information and Communication Technologies (ICT)

Can the use of ICT, both at school and outside school, support target language learning?

A variety of technological resources (directly or broadly related to ICT), with the potential to support foreign language learning, has been developed in the past decades (Garrett, 2009[15]; Butler, Someya and Fukuhara, 2014[16]; Garton, Copland and Burns, 2011[17]; Al-Mahrooqi and Troudi, 2014[18]; European Commission, 2014[19]; Laakkonen, 2011[20]). Examples range from computer labs to instructional software, social media, pedagogical and leisurely games, and so on. These resources have had an impact on teaching in the classroom, on the school environment, and on learning outside school.

Across TALIS (the OECD Teaching and Learning International Survey) countries and economies, foreign language teachers tended to make more use of ICT tools than other teachers in their lessons, even before the Covid-19 crisis (OECD, 2020[21]). This crisis later made ICT, at least in some schools, the main instrument for delivering education. This will change over the next years, as the health crisis will be overcome. However, some of the ICT-related methods and tools that became common use will probably continue to be used by teachers in the future. Therefore, it is essential to include the role and use of ICT for target language learning in this context questionnaires framework.

Elements related to ICT are included in different constructs throughout this chapter. For example, the availability of ICT tools for teaching at school is included in construct (8); the question whether teachers receive training on their utilisation is included in construct (27); the actual use of ICT tools for teaching is discussed in construct (40) and (for group and collective learning activities) in construct (35). As another example, student exposure to the target language outside school through “new media” (social media, web platforms, streaming services, computer games, language learning apps, etc.[4]) is included in construct (14).

Use of the target language for instruction in other subjects

Is student proficiency in the target language, and in other subjects, affected by attending programmes in which the target language is used for instruction in non-language-related subjects?

In many school settings, students are taught all subjects (e.g. mathematics, history or science) in one language (the language of instruction), except for one or a few foreign language classes that are taught using (at least partly) the foreign language itself. In contrast, some education programmes use more than one language for teaching non-language-related subjects. For example, in a programme in which Arabic is the language of instruction, mathematics may be taught in English; in a Spanish education programme, students may learn history in French. Such programmes can fall into different categories according to their specific characteristics for example Content and Language Integrated Learning, bilingual programmes, and other types (e.g. content-based language instruction, integration of content and language,

theme-based language teaching, and content-infused language teaching (Lasagabaster, 2008[22]; Nikula, 2016[23])). Different programmes may define and implement these categories in different ways.

Programmes in which the target language is used as a medium of instruction for other subjects can have a direct impact on students' proficiency by increasing students' exposure to the target language and giving them rich contexts in which to practice it. In addition, programmes in which multiple languages are used for instruction are thought to instil an international mindset in students; boost motivation to learn other foreign languages; and foster implicit and incidental learning by focusing on meaning and communication (Lasagabaster, 2008[22]; Cambridge Assessment, 2017[24]; Mehisto, 2012[25]; European Commission, 2014[19]).

However, some researchers have also warned on potential negative effects on equity of educational outcomes of programmes in which foreign languages are used to teach other subjects, because students enjoying more family support are more likely to enrol in and complete these programmes (Bruton, 2013[26]; Pérez Cañado, 2016[27]; Nikula, 2016[23]). Another concern for policy makers is that these programmes could also potentially harm student learning, as students may fail to understand or make progress with some subject-related content because of the language barriers they face (Nikula, 2016[28]; Marsh, Hau and Kong, 2000[29]). In practice, programmes in which the target language is used as a medium of instruction for other subjects are institutional settings that can be very different from each other. Their effect on learning probably depends on a number of institutional, regional and other factors (Annex 5.A.).

The use of the target language for instruction in other subjects has been included in the discussion throughout this chapter. The inclusion of questions about whether more than one language is used for instruction is included in construct (4), and construct (2) on target language teaching time can provide information on the amount of school time students spend in subjects taught in the target language. Constructs (26) and (27) include specific questions about the training of staff teaching other subjects in the target language. In addition, construct (28) covers information about the main specialisation of teachers (including teaching a content subject in the target language). Construct (39) encompasses whether teachers purposely integrate the learning of non-language-related content (e.g. mathematics or history) and the learning of the target language.

Government and school policies

Government and school policies frame the environment in which students learn. This policy domain contains the constructs related to the general setting for target language teaching and the school environment. Measuring the constructs in this domain would help understand the political context of the educational institutions in an education system, and its relationship with students' target language proficiency.

The setting for target language learning at school

What is the general framing for target language learning at school and how does it relate to students' proficiency?

This section includes four constructs related to the age at which students start learning the target language at school, the hours dedicated to learning, the size of target language classes, the language of instruction, and what other languages, if any, students have studied at school (Table 5.2).

The general framing for the teaching and learning of the target language, and for other languages of instruction, is often determined by a number of system- and school-level policies, guidelines and practices. The constructs included in this section are among those over which governments have the highest degree of control. Therefore, it is important for policy makers to understand the constructs most strongly associated

with students' language proficiency. These constructs also help map the overall regulatory chapter in countries/economies participating in the PISA foreign language assessment.

In addition to the information collected through the student, teacher, school and parent questionnaire, the system-level questionnaires could include specific questions on differences in the teaching of listening, reading, writing and speaking skills, e.g. about the grades in which the teaching of different skills are introduced.

Table 5.2. List of constructs: The setting for target language learning at school

Construct	Policy relevance	Level of measurement
(1) Onset of target language learning at school	Essential	Student; system
(2) Intensity of target language learning at school	Essential	Student; teacher; school; system
(3) Target language class size	Relevant	Teacher; school; system
(4) Languages learned at school	Essential	Student; school; system (availability and supply of foreign languages)

(1) Onset of target language learning at school

There is great variation among countries in terms of the age at which students begin learning foreign languages in school. The differences between countries reflect different priorities between subjects but also different considerations about what starting age is the ideal time to maximise learning opportunities and finding the right balance with the resources invested. One of the key questions for the PISA foreign language study will be to map out these differences between countries and investigate their impact on learning outcomes.

Learning onset is widely thought to influence students' proficiency in the target language, although the nature of this effect is debated. The early literature on this topic suggested that it is better for students to start learning foreign languages as early as possible (Hartshorne, Tenenbaum and Pinker, 2018[30]). Across European countries, earlier onset of foreign language learning at school is generally associated with higher proficiency in foreign languages, particularly for writing (European Commission, 2012[31]). Wilden and Porsch (2016[32]) found that early onset had a positive impact (mediated by German reading skills) on reading and listening skills in foreign languages for a sample of German students.

However, some of the literature argues that, while the "as early as possible" model fits some forms of language learning (e.g. improving immigrants' language learning in a new linguistic environment), it may not always apply to students' learning a foreign language in their native language environment. Research on Catalan schools suggests that students who started learning foreign languages later acquire communication skills faster than earlier starters for a given amount of study time (Muñoz, 2006[33]). Pfenninger and Singleton (2017[34]) found similar results in Switzerland, and suggest that the "ideal" starting age depends on contextual factors and individual competencies, attitudes and socio-emotional factors. The role of contextual and individual factors in early foreign language acquisition has been stressed by several authors (Murphy, 2014[35]; Gaonac'h and Macaire, 2019[36]). Children's development of linguistic competencies in their own languages, for example, influences their ability to acquire associated foreign language competencies (Murphy, 2014[35]). Analysing the relationship between the onset of target language learning at school and student proficiency could provide important evidence to inform national policies and priorities.

(2) Intensity of target language learning at school

"Intensity of target language learning at school" relates to the amount of in-class time allocated to target language learning and (when applicable) to learning other subjects in the target language. This can also be expressed as a proportion of students' total classroom time (across all subjects). A strong correlation between the intensity of learning at school and proficiency has been found since the earliest international surveys on foreign language proficiency (Carroll, 1975[12]; Lewis and Massad, 1975[7]). In other subjects, spending more time in classroom lessons is associated with higher proficiency (see OECD (2016[37]) for science). The relationship between classroom time and target language proficiency is likely to be mediated by several factors related to the quality of learning input and teaching, including teaching practices, availability of teachers and student attendance. In addition, Larson-Hall (2008[38]) found that for a sample of Japanese students the intensity of learning mediated the relationship between the onset of target language learning at school and foreign language proficiency.

It is important to collect information on the intensity of target language learning at school for the current school year and, in principle, also for the students' previous years in education. Such detailed information could be obtained through system-level questionnaires, but it would be more difficult to elicit from individual respondents. In that case, simpler questions (e.g. whether the amount of target language learning time has changed in recent years) could be included in student questionnaires.

(3) Target language class size

"Target language class size" is the number of students attending a typical target language class. In general, there is no evidence of a robust relationship between class size and student learning across countries (OECD, 2016[37]). However, some studies argue that smaller classes could help teaching and learning foreign languages (Aoumeur, 2017[39]). This can be justified by the role of student participation in certain foreign language teaching approaches (in particular, the communicative approach discussed in the section *Teaching approaches and methods).* For example, it could be easier to introduce learner-centred teaching or use pair work and group work in smaller classes (Garton, Copland and Burns, 2011[17]). Therefore, it is important to track, at the system level, whether countries have introduced policies specifically to reduce the size of target language classes, for example to stimulate student interaction.

While PISA already collects information on class size, it could be useful to collect this information specifically for target language classes, as it may differ from the size of classes in other subjects. On average across 48 countries and economies participating to the OECD Teaching and Learning International Survey (TALIS) survey, lower secondary teachers reported that there were 24 students in a typical ("target") class they taught (OECD, 2019[40]) (Table I.3.78). Foreign language teachers reported a slightly smaller class size (23 – the difference between foreign language and other teachers is statistically significant). In particular, foreign language teachers were more likely to teach small classes with less than 20 students. This was reported by 37% of foreign language teachers, as compared to 28% of other teachers (see Annex 5.B. for the methodology underlying this estimation).

(4) Languages learned at school

"Languages learned at school" are all the languages that the student has learned or is learning at school, whether the target language, other modern foreign languages, or ancestral and ancient languages (e.g. Latin, Greek or indigenous languages that are not widely spoken). Languages learned at school include:

- languages studied as a subject (e.g. a few hours a week are devoted to study Japanese as a foreign language – the most common approach in most countries)
- languages used to teach other subjects (e.g. English is used as a medium of instruction in history classes in a German school).. When the target language

is used to teach other subjects, it is particularly relevant to know which subjects (e.g. science or history).

Which languages students learn depends on their (or their parents') preferences, but also on curriculum prescription and the availability of language classes in schools (construct (5)). Learning additional languages can affect target language proficiency by teaching students general strategies that can be applied to the target language, for example, negotiate meaning with others, or use context for understanding. It may be associated with other general language competencies and attitudes (e.g. "multicompetence" (Rothman, Cabrelli and De Bot, 2013[41]); "multilingual competence"; "language awareness") that are valued in a comprehensive approach to language learning (Kelly, 2019[42]; Council of the European Union, 2019[43]). The number of modern and ancient foreign languages learned is positively associated with foreign language proficiency test scores across European countries, especially for reading and writing (European Commission, 2012[31]).

The school environment

What makes schools effective in fostering target language learning?

This section discusses what schools have to offer to students learning the target language and (for some constructs) other foreign languages. It covers the availability of foreign language courses, enrichment activities, and resources for teaching and learning (Table 5.3). These constructs are of direct interest to policy makers and educators. School management determines these constructs within the constraints posed by the availability of resources and by government regulations. Governments, in turn, can influence schools (especially public schools) through a variety of regulatory, organisational and financial levers. As compared to the other questionnaires, the system-level questionnaire could include more detailed questions on the regulatory environment, for example about who decides which languages are a compulsory part of the curriculum.

Table 5.3. List of constructs: The school environment

Construct	Policy relevance	Level of measurement
(5) Availability of foreign languages	Essential	School; system (policies and guidelines)
(6) School enrichment activities for target language learning	Essential	Student; teacher (participation); school (availability); system (policies and guidelines)
(7) Target language remedial lessons at school	Essential[5]	Student (participation, reason for attending); school (availability)
(8) School resources for target language teaching	Essential	School; system

(5) Availability of foreign languages

"Availability of foreign languages" refers to the choice of foreign languages available to students (independently on the languages they actually study, which is discussed in construct (4)). Student choice is constrained by which languages are offered for study, and which are compulsory for students to learn (both elements should be measured). At the school level, a wider availability of foreign languages could signal efforts to emphasise foreign language learning, for example, if the school has a specialist foreign language profile. This could have a general influence on motivation and attitudes of students, teachers and parents at the school.

(6) School enrichment activities for target language learning

"School enrichment activities for target language learning" refers to a range of extracurricular activities that can be organised within the school environment to stimulate interest in the target language and target language learning. These include, for example, target language competitions, debate clubs and

simulations (e.g. model United Nations), enrichment lessons, visits to the school by students from other language communities, and various projects related to target language learning. They could also include setting up environments where communication with students outside the classroom happens in a foreign language (for example, communication with teachers or in a canteen, direction signs). Those opportunities can facilitate target language learning by improving the "learning conditions" (Lightbown and Spada, 2013[44]) of students in the school. Across education systems that participated in the PISA 2015 survey (OECD, 2016[37]), students in schools offering science competitions were, on average, more proficient in science and more likely to expect to work in a science-related occupation than other students.

(7) Target language remedial lessons at school

"Target language remedial lessons at school" refers to opportunities offered by schools to low-performing students for additional lessons in the target language. Students may attend them voluntarily or because they are required to do so, which could affect their motivation (Carroll, 1963[45]). The association between attending remedial lessons and target language proficiency is expected to be negative at the student level, as students attend them because they are less proficient (in contrast to enrichment activities; construct (6)). At the school and system levels, the availability of remedial lessons could reduce the number of low-performing students.

This construct would be considerably more useful for policy analysis if it measured not only for the current year, but also for previous years (how long a system of target language remedial lessons has been in place in a school; or how long a student has attended remedial lessons).

(8) School resources for target language teaching

"School resources for target language teaching" refers to the availability of financial resources (e.g. budget for student exchange visits), goods (e.g. target language books, DVDs) and ICT tools (both in terms of physical infrastructure, e.g. computer labs, and software related to foreign language learning, teaching and assessment (Garrett, 2009[15])). This can refer to actual availability, but also to identified constraints. The available resources can support teaching and school activities, provided that teachers make effective use of the materials available at the school (see constructs (40) and (41)).

The availability of ICT resources influences the opportunity to learn the target language through ICT, and is associated with students' proficiency in foreign languages across European countries (European Commission, 2012[31]). The availability of school resources for target language teaching is also of interest because it can be related to equitable opportunities to learn the target language. For example, across education systems that participated in the PISA 2015 survey, students from lower socio-economic backgrounds tended to be in schools with less educational resources (OECD, 2016[46]).

Students and learning

Students' support, motivation, opportunities and ability to learn foreign languages are influenced by a large variety of factors. This policy domain includes the constructs related to students' background, environment and family support; their attitudes, motivations and behaviours; and the relationship between target language learning and intercultural and multilingual environments.

Student background, environment and family support

How is target language learning affected by students' background and their experiences outside of school?

Students' proficiency in the target language depends not only on their learning at school, but also on their background and environment. Some students may speak the target language at home with their family, or be in an environment where they have other opportunities to use it, thereby improving their proficiency.

Opportunities to travel abroad are also very different across countries, schools and families. In addition, families may offer support with homework, or simply by transmitting to their children positive attitudes and motivations towards target language learning.

Some of the constructs discussed in this section (Table 5.4) are of direct policy relevance as they can be changed by governments or schools. For example, student visits to other language communities can be organised by schools. Other constructs are more important to inform students and parents, such as parents' support for target language learning. In addition, some constructs provides essentially contextual information for comparing proficiency data across students and countries (e.g. language background).

Table 5.4. List of constructs: Student background, environment and family support

Construct	Policy relevance	Level of measurement
(9) Language background	Essential	Student; parent
(10) Parents' target language proficiency	Relevant	Parent
(11) Family support in target language learning	Relevant	Student; parent
(12) Family and peers' perceptions and attitudes related to the target language	Relevant	Parent; teacher
(13) Family and peers' perceptions and attitudes related to target language lessons	Relevant	Parent; teacher
(14) Target language exposure through the media	Essential	Student, parent (actual exposure); system (policies and widespread practices)
(15) Student visits to other language communities	Relevant	Student, parent (actual experiences); system (opportunities)
(16) Face-to-face exposure to and use of target and foreign languages outside of school	Essential	Student; parent

(9) Language background

"Language background" is the set of languages used by the student at home and with closely related individuals, particularly family members. It is important to know which languages the student has been exposed to at home and how proficient he or she is in each language (and skills, e.g. listening, speaking). Measures of students' language background would also allow for identifying the diversity of languages spoken in a school.

Language background can influence students' target language proficiency in at least three ways. First, some students may speak the target language at home, with a direct effect on their level of proficiency.

Second, learning different languages during childhood provides metalinguistic insights that change the way children think about language (Barac and Bialystok, 2010[47]). Multilingualism can enhance the acquisition of additional languages, especially if learners possess a high level of literacy in the languages they have used since childhood (Cenoz, 2003[48]). In addition, multilingual children could have more positive attitudes than other children towards language learning in general (Brown, 2007[49]).

Third, the language(s) spoken since childhood can influence target language learning (Brown, 2009[50]), and even interfere with it (Amirabadi and Razmjoo, 2017[51]; Derakhshan and Karimi, 2015[52]; Brown, 2009[50]), depending on the distance amongst these languages (Derakhshan and Karimi, 2015[52]; Lindgren and Muñoz, 2013[53]) (for more information on language distance, see box 4.1 "The effect of L1 and FL distance on FL learning" in chapter 4). Therefore, language background can also be looked at in conjunction with the information provided by indices of distance between languages available in the literature on comparative linguistics (Isphording and Otten, 2014[54]; Lindgren and Muñoz, 2013[53]).

(10) Parents' target language proficiency

"Parents' target language proficiency" is related to the parents' language skills and mastery of the target language. This construct can be measured, for example, through parents' self-assessed proficiency or the reported difficulty in performing some tasks[6] in the target language (e.g. reading a newspaper article); and through proxy factors (e.g. use of the target language at work, the target language teaching onset for parents). The available evidence suggests that parents' target language proficiency is positively associated with higher proficiency in foreign languages amongst students in Japan (Yoshitomi, 1990[55]), but also across European countries (European Commission, 2012[31]; Bonnet, 2002[56]; Lindgren and Muñoz, 2013[53]). Being proficient in the target language can make it easier for parents to help children with homework and with target language learning through leisure activities. In addition, the proficiency of parents could change students' perceptions of and attitudes towards the target language (Yoshitomi, 1990[55]). For example, proficient parents could be role models for their children, indicating that learning the target language is realistic and useful. Proficient parents may also be able to offer their children more exposure to the target language.

(11) Family support in target language learning

"Family support in target language learning" refers to the help given (by family members) and received (by the student) at home specifically for target language learning. This support could include help with homework and preparing for tests, and practicing the target language together. Students can benefit from support from parents and siblings, but also from the extended family (e.g. aunts and uncles) (Cumming, 2012[57]).

Families play an important role in fostering academic and non-academic success for students (OECD, 2019[4]), for example by developing students' self-confidence. Parental involvement in homework does not seem to be robustly associated with student attainment (possibly because it is associated with lower student achievement, or because of limited content or pedagogical knowledge), but it still fosters positive learning-related attitudes, ideas and behaviours in students (Hoover-Dempsey et al., 2001[58]). The effect of family support on target language proficiency is also likely to depend on other factors included in this chapter, for example language background and parental target language proficiency.

(12) Family and peers' perceptions and attitudes related to the target language

"Family and peers' perceptions and attitudes related to the target language" refers to perceptions and attitudes that can directly or indirectly affect students" own perceptions, with a potential effect on target language learning (Bartram, 2018[59]). It includes the perceived difficulty in learning the target language; and the perception of usefulness (or uselessness) of the target language for young people.

(13) Family and peers' perceptions and attitudes related to target language lessons

"Family and peers' and attitudes related to target language lessons" refers to a variety of feelings, attitudes and opinions with respect to target language lessons at school. It includes the degree of satisfaction with the target language teacher and the teaching methodology, and their perceived effectiveness in improving students' proficiency. Parents' perceptions and attitudes can directly or indirectly affect students' own perceptions, either positively or negatively. Family, friends and peers are all major influences on students' attitudes towards language learning (Bartram, 2018[59]).

(14) Target language exposure through the media

"Target language exposure through the media" refers to the way and amount of time students consume media in the target language outside of school, either with the purpose of learning or for leisure. "Media" refers to both "traditional media" (television, books, radio, magazines, etc.) and "new media" (social media, web platforms, streaming services, computer games, language learning apps, etc.). Media content can be

presented in the target language in different ways (e.g. it can be dubbed or subtitled in the target or local language).

Questions could distinguish between different media channels. This would allow for differentiating amongst media-related activities that provide opportunities to develop passive or active language use: target language listening (e.g. listening to songs or watching a video series); speaking (e.g. engaging in some types of games or in video-calls); reading (e.g. reading books or blogs); writing (e.g. e-mails, chatting on line); or interaction (which could involve several of the activities listed above). It would also allow for investigating the relative benefits of visual and aural content (for example, Kim and Kim (2011[60]) find that upper secondary Korean students have a visually oriented learning style, well-suited to visual learning aids, such as books and videos).

Across European countries, there is a strong positive association between students' target language exposure through traditional and new media, and target language proficiency (European Commission, 2012[31]). Researchers have found a positive association between using English language games and scores in English proficiency tests amongst children of different ages in Japan (Butler, Someya and Fukuhara, 2014[16]) and Sweden (Sylvén and Sundqvist, 2012[61]). Exposure to media could influence language learning even unconsciously, as can exposure to music in the target language (Grant, 2012[62]). In addition, exposure through the media could be influenced by the students' target language proficiency, as more proficient students would find it easier to consume media in the target language.

Exposure to and use of the target language through various media can be influenced by both personal or social habits (e.g. a disposition towards enjoying cultural content in the original language, without dubbing) and policies and practices by private and public bodies (e.g. a specific policy to broadcast movies in the target language or with target language subtitles on public television channels).

(15) Students' visits to other language communities

"Students' visits to other language communities" refers to experiences for students' travel to other language communities where they can practice the target language. Collaborating and interacting with other speakers are fundamental elements of foreign language learning processes (Donato, 1994[63]; Lantolf, 2000[64]; Lightbown and Spada, 2013[44]). These could include holiday trips, family visits, short-time immersion programmes and school trips (including school exchange visits). Information should be collected on the type of trip, whether the target language was used, and on the length and frequency of the trips, as the simple fact of travelling to another linguistic community does not ensure an opportunity to practice the target language. Visits to other language communities can provide an opportunity to practice the target language even if it is not the community language, as long as the target language is used for communication as a *lingua franca* (for example, English is widely used by tourists and exchange students in non-English-speaking countries).

(16) Face-to-face exposure to and use of target and foreign languages outside of school

"Face-to-face exposure to and use of target and foreign languages outside of school" refers to use of opportunities to interact in person with other people in the target language (excluding visits from or to other language areas). This can include the frequency of talking with tourists, speaking face-to-face with friends and family, or interacting with peers in the target language. Many people encounter foreign languages in their everyday surroundings; for example, in 2007, around 90% of Finns reported hearing foreign languages in their environment, with English as the most commonly heard language (Leppänen et al., 2011[65]). This could increase motivation (construct (17)), by helping learners imagine themselves as target language users. In addition, face-to-face interactions in the target language provide opportunities to improve listening and speaking skills.

Learners' attitudes, motivations and behaviours

What are students' attitudes, motivation and behaviours towards the target language and target language teaching, and how are they related to proficiency?

Attitudes, motivations and behaviours (Table 5.5) related to the target language function as a dynamic system, interacting with each other over time and in complex ways (Dörnyei, 2010[66]). They depend on a variety of factors, from the broad social and economic context, to individual cognitive abilities, possibilities and potentials, as well as to the process of learning itself (Dörnyei, 2005[67]; Mercer, 2011[68]). Many of these factors are outside the control of schools and national governments. Nonetheless, they are the context in which learning takes place, and must be taken into account when comparing proficiency across countries.

Attitudes, motivations and behaviours, however, can also be modified by education policies and teaching practices. They can even be recognised as an outcome of the education process itself, as stimulating interest and curiosity can be considered a goal of education. For example, interest in the presence of foreign languages in one's daily life and in their speakers' socio-cultural world is an official learning target in the Flemish Community in Belgium (see e.g. Flemish Ministry of Education and Training (2019[69]) for primary education). This makes the relationship between teaching practices, school and national policies on the one hand, and learners' attitudes and motivation, on the other, an interesting policy question in itself.

Table 5.5. List of constructs: Learners' attitudes, motivations and behaviours

Construct	Policy relevance	Level of measurement
(17) Students' motivation for target language learning	Essential	Student
(18) Students' perceived level of proficiency	Relevant	Student
(19) Students' attitudes towards target language learning at school	Essential	Student
(20) Time spent on target language study, homework and other structured learning activities outside of school	Essential	Student (self-reported time); parent (observed time); teacher (recommended or ideal time, frequency of assigning homework)

(17) Students' motivation for target language learning

In foreign language acquisition, "motivation" refers to "the extent to which the individual works or strives to learn the language because of a desire to do so and the satisfaction experienced in this activity" (Gardner, 1985[70]); it is one of the main correlates of student language learning (Dörnyei and Skehan, 2003[71]). In recent years, interest has grown in measuring skill-specific motivation (see for example Lee, Yu and Liu (2018[72]) for writing), which could be of particular interest.

Motivation includes extrinsic and intrinsic orientations. Intrinsic orientations are reasons to learn the target language that are connected with the inherent pleasure and interest in learning the language, for example because of the satisfaction of individual curiosity or the enjoyment of learning the language (which could be related to teaching approaches and methods). Extrinsic orientations are reasons that are instrumental to consequences, such as earning high grades, working in a stimulating career, being admitted to prestigious universities, or travelling across cultural boundaries (Noels et al., 2003[73]; Melzi and Schick, 2012[74]).

Based on the "L2 motivational self system" chapter (where L2 means second language learning), items on what students would like to become in the future also seem important for capturing motivation. This chapter identifies three primary sources of motivation to learn a language (Dörnyei and Chan, 2013, pp. 438-439[75]): the ideal L2 self (i.e. "the person we would like to become"), stimulating an internal desire to learn; the ought-to L2 self, representing social pressures (e.g. finding a job, avoiding parental disappointment) to

learn the foreign language; and the actual experience of learning, which can be more or less or less enjoyable.

Studies from China, Hungary, Iran and Japan show that the ideal self is a stronger predictor of motivation than the ought-to self (Taguchi, Magid and Papi, 2009[76]). In China, the ought-to self predicts motivation better than in other countries (Taguchi, Magid and Papi, 2009[76]), but a study of young learners in Hong Kong still found no correlation between ought-to self measures and English and Mandarin grades (Dörnyei and Chan, 2013[75]).

(18) Students' perceived level of proficiency

"Students' perceived level of proficiency" is a self-assessment of students' own level of proficiency in the various language skills in the target language. This construct should be based on the same definitions of proficiency used for the PISA 2025 cognitive proficiency assessment. Students can assess their own proficiency by stating what they can do through a self-assessment tool (e.g. following or giving a talk on a familiar topic (Council of Europe, 2001[1]; 2018[77])) or by providing a global self-assessment, both of which have been found to be correlated with actual proficiency (Berns, de Bot and Hasebrink, 2007[78]). However, the ability of students to assess their own proficiency cannot be assumed, and it varies across genders and countries (Denies and Janssen, 2016[79]). The perceived level of proficiency could affect students' attitudes and motivation towards learning the target language. Results from the IEA Six-subjects study in the 1970s found an association at the student level between low self-perceived English proficiency and time spent learning English relative to other subjects (Lewis and Massad, 1975[7]).

The ability to self-assess can be considered a learning goal in itself as it is important to stimulate self-directed and lifelong learning (Little, 2005[80]; Denies and Janssen, 2016[79]). This ability can be proxied by the distance between self-assessed proficiency and the actual proficiency measured through the PISA foreign language assessment.

(19) Students' attitudes towards target language learning at school

"Students' attitudes towards target language learning at school" includes subject-specific anxiety and self-concept, which have been shown to be strongly related to student proficiency (OECD, 2016[46]) (*PISA 2022 Assessment and Evaluation Chapter*, forthcoming). They can also include the perceived difficulty (European Commission, 2012[31]) of target language learning, which is negatively associated with proficiency across European countries (European Commission, 2012[31]). Students' attitudes can be assessed relative to those of other students (e.g. perceived difficulty), but also to those regarding other subjects.

Given that PISA measures subject-specific attitudes for other subjects (self-concept, anxiety), it is advisable to measure attitudes towards the target language using scales that are as close as possible to those used for the other subjects. In addition, given that proficiency is measured separately for each communication skill, it could be useful to measure at least one of these subject-specific attitudes for each of these skill (e.g. specifically for speaking, or for reading).

(20) Time spent on target language study, homework and other structured learning activities outside of school

"Time spent on target language study, homework and other structured learning activities outside of school" refers to time spent on study, homework and structured target language learning activities outside the classroom, and to the regularity of these activities. This includes ordinary homework and study, collaborative assignments, as well as preparation for tests and assessments or group work on assignments. Time spent learning the target language through other structured learning activities (e.g. private tutoring or target language learning camps) should also be included, separately from homework and classroom-related study.

Study and homework can supplement classroom activities, and existing studies find a modest positive association with student academic achievement (Cooper, Robinson and Patall, 2006[81]), even though this association is less strong for Asian countries (Fan et al., 2017[82]). However, the direction of the empirical association between this construct and target language proficiency is not clear *a priori*. Students struggling to reach the level of proficiency required in their class may study more, resulting in a negative association between time spent on study and homework, and proficiency. A negative association between proficiency and time spent studying outside of school has been observed for science, for example OECD (2016[37]).

Other target language learning activities, such as private tutoring, can also be undertaken by low-performing students to catch up with their peers. However, such activities are increasingly offered to students as supplementary learning opportunities by parents who can afford it (see Nunan (2003[83]) for the case of the Asia-Pacific region; and Cronquist and Fiszbein, (2017[84]), for Latin America). Therefore, it could be useful to measure not only current target language activities outside school, but also whether the student has been involved in such activities in the past.

Intercultural and multilingual environments and target language learning

What is students' understanding of intercultural and multilingual environments, and how is it related to their proficiency?

In many countries and communities, society is becoming increasingly diverse, for example through the influx of immigrants and the rise of new, complex forms of citizenship and belonging (OECD, 2019[85]). The expansion of mobility is making the world increasingly multilingual (Jenkins, 2017[86]; King, 2017[87]). Cultural awareness and the ability to interact respectfully with people from different backgrounds can help diverse cultures live peacefully in close proximity and find solutions to common problems. This increase in diversity is prompting policy makers and educators to find ways to teach young people how to challenge biases and stereotypes towards other cultural and language backgrounds through intercultural dialogue (OECD, 2019[85]; Council of Europe, 2008[88]).

Foreign language learning may be related to the understanding and appreciation of intercultural and multilingual environments in a variety of ways (Table 5.6). Students' appreciation of cultural and linguistic diversity may be enhanced by the study of a foreign language, both through the understanding of the challenges and benefits of speaking different languages and through exposure to foreign language content (e.g. literature, news) related to other cultures. In addition, a deeper appreciation of cultural and linguistic diversity can motivate students to learn foreign languages (Della Chiesa, 2012[89]). Therefore, intercultural and multilingual competencies are both an input into and an outcome of the process of language learning (an example was given in the section *Learners' attitudes, motivations and behaviours*: interest in foreign language speakers' socio-cultural world is a learning target in the Flemish Community of Belgium).[7] In addition, languages enjoying the status of a *lingua franca* (whether English or other languages (Pütz, 1997[90])) are often used for intercultural communication. Negotiating meaning between people from different cultures therefore becomes an essential component of learning and using the language (Jenkins, 2017[86]; Seidlhofer, 2011[91]).

Table 5.6. List of constructs: Intercultural and multilingual environments and target language learning

Construct	Policy relevance	Level of measurement
(21) Linguistic and cultural diversity in the community	Relevant	Student; parent; system
(22) Respect and openness towards people from other culture and language backgrounds	Relevant	Student
(23) Pluricultural and plurilingual education	Relevant	Student; teacher; school; system

(21) Linguistic and cultural diversity in the community

"Linguistic and cultural diversity in the community" relates to the variety of languages and cultures that students can experience within their communities. It is a characteristic of the community (not of the student, in contrast to constructs in the section *Student background, environment and family support*). This construct includes the proportion of people who speak a language different from the local language at home or who come from other countries, as well as the number of languages and dialects spoken in a certain region. It can also include information on relationships with target or foreign language-speaking countries (e.g. if there are many foreign tourists, or if multilingualism is encouraged as a national or regional policy). Being exposed to a multilingual or multicultural environment can change students' attitudes towards and perceptions of foreign languages, positively influencing their motivation to learn the language (Burstall, 1975[92]; Norton and Toohey, 2001[93]; Lightbown and Spada, 2013[44]).

(22) Respect and openness towards people from other culture and language backgrounds

"Respect and openness towards people from other culture and language backgrounds" involves sensitivity towards, curiosity about and willingness to engage with other people and other perspectives on the world ("openness"); and positive regard and esteem for cultural and linguistic differences based on the judgement that they have intrinsic importance, worth or value ("respect") (adapted from OECD (2019, p. 175[94])). This construct is based on the concept of respect and openness to other cultures described in the PISA Global Competence chapter (OECD, 2019[94]). This construct should emphasise the willingness to engage with interlocutors who are not only culturally, but also linguistically different (see, for example, "facilitating pluricultural space" (Council of Europe, 2018, p. 122[77]), and emphasise interactions within the students' environment (rather than a mere interest in "exotic" experiences (OECD, 2019[94])).

Knowing more than one language is arguably related to dispositions towards other cultures (OECD, 2019[94]). Openness towards dissimilar others and a willingness to approach them are parts of "international posture", a concept positively associated with foreign language learning motivation and proficiency (Yashima, 2002[95]; 2013[96]). Respect and openness towards people from other cultures and language backgrounds could both influence students' target language learning and be influenced by it (for example, if they are integrated in foreign language learning at school; see construct (23)).

(23) Pluricultural and plurilingual education

"Pluricultural and plurilingual education" refers to educational activities in school and in the classroom to educate students about cultural and language diversity, and especially about the diversity that can be experienced in the students' communities. This diversity includes, for example, the presence of multiple languages, dialects, religions and lifestyles. School and classroom activities can include:

- Activities to develop students' plurilingual and pluricultural communication capabilities (Heugh, 2018[97]) in the classroom, for example:
    - through their inclusion amongst the learning goals of the teaching of the target language or other subjects (European Commission, 2015[98])
    - through teaching the concept of *lingua franca* and its role for intercultural and global communication (Seidlhofer, 2011[91]; Graddol, 2006[99])
- promoting initiatives to learn about the traditions of different cultural groups or pluricultural events at school
- encouraging or creating opportunities for multilingual students to make use of their full linguistic repertoire in the school environment (e.g. through heritage or mother tongue-language teaching (Cummins, 2005[100]; European Commission, 2015[98])).

Pluricultural and plurilingual education have been part of foreign language learning theory and practice for a very long time (Ollivier, 2019[101]). They could be related to students' target language learning by increasing students' respect and openness towards people from other cultures and language backgrounds (see previous section). In addition, this construct is also related to "linguistic and cultural diversity in the community", as schools in culturally diverse environments need to encourage intercultural sensitivity and help students move away from ethnocentric world views towards tolerance, acceptance, respect and appreciation of other cultures (OECD, 2019[94]).

Teachers' training and profile

Teachers play a fundamental role in learning. Their training and hiring are major investments for governments, making it essential to understand what makes an effective foreign language teacher. Around 18% of teachers around the world are foreign language teachers. Foreign language teachers have a different profile from others: for example, they are better prepared for teaching in multicultural contexts and more likely to have studied abroad (OECD, 2020[21]).

This policy domain is related to human resource policies (and teacher characteristics). Consistent with the scope of this chapter, this domain covers domain-specific constructs tied to foreign language learning. Therefore, some important constructs related to teaching in general (e.g. collaborative practices (Ainley and Carstens, 2018[5])) are not discussed.

An intuitive choice to measure constructs in this domain is often through the teacher questionnaire. However, an important analytical constraint of this approach is that it will not be possible to link responses to the teacher questionnaire with information on student proficiency (limiting the analysis to associations at the school or system level). This consideration will have to be kept in mind (together with other technical considerations) when choosing how to measure these constructs through a combination of different levels of measurement (student, teacher, school, system).

Human resources

How do policies on teaching staff hiring and training affect target language learning?

Governments and schools shape the teaching force through the regulations, incentives, policies and practices they put in place. Different requirements and expectations can apply to the training of target language teachers, their recruitment and their subject specialisation. Incentives of various types can be provided to hire and retain staff with different characteristics (e.g. teaching assistants or more-experienced teachers). These policies can influence teaching effectiveness and therefore student learning. They also affect the attractiveness and accessibility of the profession to potential teachers and, in turn, the availability (or scarcity) of teaching staff.

Human resource policies must fit the broad context and the teaching approaches used in a particular country or school. For example, in various historical periods Korea and Japan invested massively in teacher training to tackle shortages of foreign language teachers (Chang, 2012[102]), while other educational systems may have more than enough candidates for the available jobs. In addition, some specific forms of training may be required for teachers using the target language to teach other subjects.

The system-level questionnaire could collect information on human resources in primary to upper secondary education, as all these levels could affect the cumulative learning of 15-year-old students. In addition, the system-level questionnaire could ask whether there are differences in human resource policies related to foreign language and other teachers. This information is related to several of the constructs presented in this section (Table 5.7), and it would provide a broader context for their interpretation.

Table 5.7. List of constructs: Human resources

Construct	Policy relevance	Level of measurement
(24) Target language teaching experience	Essential	Teacher
(25) Target language teacher and staff availability	Essential	School; system
(26) Teachers' initial education and qualifications	Essential	Teacher (actual training received); school, system (guidelines and regulations)
(27) Teachers' in-service training	Essential	Teacher (actual training received); school, system (guidelines and regulations)
(28) Teaching specialisation	Essential	Teacher (actual taught subjects); school, system (policies and guidelines)

(24) Target language teaching experience

"Target language teaching experience" refers to the number of years of experience and the type of experience that target language teachers have. This includes the number of years teaching the target language, both in general and to the age group that is currently being taught. Teaching experience could be positively associated with student language learning, for example if teachers learn how to adapt to different classes and address common linguistic mistakes made by students. However, the association could also be negative if, for example, more-experienced teachers are slower in adopting new teaching practices.

A moderately positive association between teacher experience and student achievement has been found in the United States across different disciplines (Hanushek and Rivkin, 2004[103]; Leigh, 2010[104]). Across European countries, the duration of teachers' placement in the same school or the number of languages they taught in previous years is not strongly associated with students' target language proficiency. However, teachers' experience in target language teaching was positively associated with proficiency in the SurveyLang study (European Commission, 2012[31]).

(25) Target language teacher and staff availability

"Target language teacher and staff availability" indicates the extent to which target language teaching personnel is available for covering the planned long-term needs, temporary vacancies and short term replacements. Teaching staff shortage can lead to employing unqualified staff for teaching the target language, with a negative impact on learning.

Teaching staff can include:

- teachers (i.e. personnel with the required qualifications to teach the target language)
- auxiliary staff (i.e. teaching assistants) with high proficiency in the target language. This type of staff is widely used, for example, in the implementation of Content and Language Integrated Learning (CLIL) in Spain, but they could be used to help teachers in any programme. Highly proficient teaching assistants can improve the learning process by making target language communication in the classroom more authentic and by reacting quickly to oral or written production (Bruton, 2011[105]; Dafouz and Hibler, 2013[106])
- other auxiliary staff (e.g. technical staff for language labs).

Besides the number or presence of teaching assistants, it is also important to collect information on their role (regular school staff or staff coming through special arrangements, such as exchange and guest programmes), and on whether they received some training or induction when (or prior to) starting their assignment in the school.

(26) Teachers' initial education and qualifications

"Teachers' initial education and qualifications" refers to the education and training the target language teacher undertook to become a teacher (independent of whether the teacher is specialised in the target language or was trained as a general teacher). This encompasses degrees, post-graduate certifications and specialisations and any training required to become a target language teacher. Information could be collected on:

- level of training
  - required ISCED level for target language teachers
  - requirements related to certificates or degrees, such as a requisite for the degree and additional specialisation courses in the field of target language studies, general education or other subjects (particularly for teachers teaching other subjects in the target language; see the section *Use of the target language for instruction in other subjects* and the report by European Commission/EACEA/Eurydice (2017, p. 91[11]))
- content of the training
  - general pedagogical knowledge
  - target language proficiency
  - target language-specific teaching and assessment practices (Coombe, Troudi and Al-Hamly, 2012[107]) related to the age group being taught
  - experience in the classroom
  - for teachers involved in teaching other subjects (e.g. mathematics or science) in the target language, methods for supporting target language learning while delivering instruction in another discipline.

Across the 33 countries and economies with available data from TALIS, most lower secondary education teachers are qualified with a single credential for studies in subject-matter content (and possibly other subjects) and pedagogy (OECD (2019[40]), Table I.4.12). Overall, the same is true for foreign language teachers, but with important differences at the system level (OECD, 2020[21]). In some education systems, particularly those experiencing teacher shortages, many teachers do not complete any formal teacher education (see Cronquist and Fiszbein (2017[84]) for a discussion of this problem in Latin American countries). Among TALIS countries and economies, the share of lower secondary teachers without a formal qualification for the subjects they are teaching is particularly large (and not significantly different for foreign language and other teachers) in Saudi Arabia (10%) and Mexico (8%).

Teacher training, especially when combined with practical experience, is expected to increase teachers' effectiveness in the classroom. Previous research has shown that teachers' educational attainment (i.e. the highest level of education obtained) is not robustly related to students' foreign language proficiency across European countries, perhaps because of the limited variation amongst these countries. In contrast, research suggests that having a specialised certificate for teaching the target language is positively associated with students' language test scores (European Commission, 2012[31])

(27) Teachers' in-service training

"Teachers' in-service training" refers to continuing professional development (Cordingley et al., 2015[108]) that target language teachers undertake as part of their job, because it is required, incentivised or simply offered by the school or other organisations (e.g. the ministry or the teachers' union). It can be attended in person or on line. In-service training could play an important role for developing the skills and competencies to teach foreign languages (Garton, Copland and Burns, 2011[17]), but it remains inadequate in many countries (Prapaisit de Segovia and Hardison, 2008[109]; Nunan, 2003[83]). On average across OECD countries and economies, 82% of lower secondary education teachers reported that in-service

training and development activities (e.g. courses or seminars) had an impact on their work. Reporting a positive impact of training is highly correlated with teachers' job satisfaction and self-efficacy in most TALIS-participating countries and economies (OECD, 2019[40]). When properly designed, in-service training can also be a tool to improve the implementation of the government's education policies in the classroom (Rixon, 2017[110]).

It is important to measure whether the following elements were included in any in-service training:

- foreign language-learning pedagogy, including training aimed at improving teachers' assessment literacy (see the section *Use of the target language for instruction in other subjects*).
- improving teachers' target language proficiency.
- using teaching materials and infrastructure, including ICT, for target language teaching. For analytical purposes, the materials included in this construct should be aligned with those included in "school resources for target language teaching" (construct (8)) to understand how closely training and availability are related.
- for teachers involved in teaching other subjects (e.g. mathematics or science) in the target language, methods for supporting target language learning while delivering instruction in another discipline. It is important to investigate if their training focuses on language pathways or on the specific methodology for using a foreign language for instruction in other subjects.

Research has suggested that a didactic model in which facilitators simply tell teachers what to do, or give them materials without giving them opportunities to develop skills and inquire into their impact on pupil learning is not effective (Cordingley et al., 2015[108]). The following information could help better understand teachers' motivation and usefulness of training:

- whether teachers found the training beneficial and it can help reshape their teaching style (i.e. they can apply what they have learnt).
- whether attendance was voluntary or obligatory, and whether any incentives were linked to it (e.g. voluntary, but required or instrumental for promotions or salary increases).

Given the nature of in-service training as a form of training over teachers' careers, it is important to ask teachers not only about the training received in the current year, but also in the past few years.

(28) Teaching specialisation

"Teaching specialisation" refers to the range of subjects taught by teachers or that teachers have taught in the past, and for how long, as courses teachers teach may be determined by the needs of the schools that employ them. Teachers can be classified in one of four categories:

- teaching the target language only (specialised in the target language)
- teaching the target language and other foreign languages (specialised in foreign languages)
- teaching the target language and other subjects (different from foreign languages)
- teaching a content subject in the target language (see the section *Use of the target language for instruction in other subjects*).

In addition, teachers who are not specialised in the target language can only teach the target language as their main subject (the one on which they spend the largest portion of their working time), as a secondary subject (if there is another subject on which they spend more time), or on an equal basis with another subject (dual specialisation).

More-specialised teachers have more opportunities to understand and learn to tackle the specific challenges students face when learning the target language. Therefore, there may be a positive association between teacher specialisation and target language learning. However, within some countries there may be very little variation in teaching specialisation, as the subjects that can be taught are often regulated at the national level.

The teacher

Are teacher characteristics, attitudes and behaviours associated with student target language learning?

This section deals with the characteristics, attitudes and behaviours of teachers that could be important for foreign language learning (Table 5.8). These characteristics, attitudes and behaviours are embodied in the teachers or they depend on their personal choices. Therefore, governments and schools do not control them directly. However, the constructs in this section are highly relevant to policy makers because governments and schools can influence them in different ways, including in recruitment and training policies or by offering incentives of various types.

Table 5.8. List of constructs: The teacher

Construct	Policy relevance	Level of measurement
(29) Teachers' visits to other language communities	Relevant	Teacher; school, system (funding and incentives)
(30) Teachers' attitudes related to target language teaching	Essential	Teacher; student; school
(31) Teachers' target language proficiency	Essential	Teacher (actual proficiency); system (guidelines and expectations)

(29) Teachers' visits to other language communities

"Teachers' visits to other language communities" refers to experiences through which teachers have the opportunity to interact in the target language with people from other language communities, by teaching or collaborating with colleagues in other language communities, through tourism, training, family visits, etc.. Through this interaction, teachers can potentially develop linguistic and intercultural communicative competencies (Cuenat, Bleichenbacher and Frehner, 2016[111]).

An earlier study indicated that across European countries, there is not a robust relationship between generic stays abroad by teachers and their students' proficiency (European Commission, 2012[31]). However, stays abroad are heterogeneous and their relationship with language learning could depend on the type of activities in which teachers were engaged. Therefore, it is important to measure not only the occurrence of the visits, but also their duration and relation with the target language (e.g. full immersion in a target language country; or interacting in the target language with non-native speakers in a country with a different language). It could also be useful to know whether these visits were preceded by or followed-up with specific training or discussions with colleagues or school management.

As compared to other types of stays abroad, exchange visits include learning and sharing of teaching practices, pedagogical knowledge, and other information related to teaching in a structured way. Therefore, exchange visits and more generic stays in other language communities should both be measured as they can affect teachers' language proficiency and attitudes towards target language teaching; they can also have a direct influence on target language teaching practices. At the school and national levels, it is relevant to measure the existence of resources for teachers' travel for didactic purposes (e.g. funding related to international programmes).

(30) Teachers' attitudes related to target language teaching

"Teachers' attitudes related to target language teaching" refers to a variety of attitudes specifically related to the teaching of the target language (as distinct from perceptions originating from classroom management and the dynamics between students and teachers). This construct includes enjoyment in and perceived difficulty of teaching the target language, but also whether teachers think certain teaching approaches are effective. Teachers' attitudes towards the target language class can affect students' attitudes and motivation (Horwitz, 1990[112]; Kern, 1995[113]; Stern, 1983[114]) and therefore potentially influence language learning.

(31) Teachers' target language proficiency

"Teachers' target language proficiency" refers to teachers' language skills and mastery of the target language. Teachers' low proficiency (and possibly the associated low confidence) with the target language can hinder student learning (Butler, 2004[115]; Garton, Copland and Burns, 2011[17]; Nunan, 2003[83]) and hinder implementation of the communicative teaching approach (Kuchah, 2009[116]). For example, in the context of English teaching, low target language proficiency has been identified as one of the key barriers to English learning in Latin American countries (Cronquist and Fiszbein, 2017[84]).

However, recent research suggests that target language teachers not only need general language proficiency but a specialised subset of language skills required in the classroom context (e.g. "English for Teaching", a bounded form of English for Specific Purposes used in the classroom (Freeman et al., 2015[117])). The idea is that the language proficiency should help in managing the classroom, understanding and communicating lesson content, and assessing students and giving feedback.

Measuring target language proficiency without testing teachers is a serious challenge. A realistic approach could include:

- Generic self-assessment items, following a similar approach as that described for perceived student proficiency (construct (18)).
- Self-assessment items on using the target language in a classroom setting. Examples could include giving instruction to students on how to carry out exercises; keeping order in the classroom; and explaining grammatical issues in the target language.
- Proxies of general target language proficiency such as having earned education degrees in the target language (and posterior validation of these degrees by refresher trainings and certificates); having lived in target language-speaking communities for a long period (e.g. more than one year); or having spoken the target language since childhood. For example, in Japan and Korea programmes to attract highly proficient teachers from English-speaking communities (such as the Japan Exchange and Teaching programme in Japan and the Korea English Teacher Training Assistant programme) have played an important role in foreign language teaching (Chang, 2012[102]).

Teaching practices

This domain discusses current practices and pedagogical thinking. This policy domain is related to teaching approaches and methods; assessment practices; and the use of the target language for the instruction in other subjects.

Measuring current teaching practices is important to understand the landscape of target language teaching in an education system. However, current teaching practices are not always the same as the past practices

students have been exposed to (and past practices are difficult to recollect and measure). This could weaken the relationship between current practices and student proficiency.

TALIS provides evidence that, in many education systems, foreign languages are taught in a different way than other subjects. This evidence does not imply or suggest that foreign language teachers teach in ways that are more or less adequate, modern or effective than other teachers. However, it does provide support for the need to improve the understanding and gather domain-specific evidence of how foreign languages are taught across education systems.

For example, among the 4 cognitive activation teaching practices measured in TALIS (see OECD (2019[40]), Figure I.2.1), foreign language teachers are significantly less likely than other teachers to report to frequently "give tasks that require students to think critically" (by 10 percentage points); "ask students to decide on their own procedures for solving complex tasks" (by 8 percentage points); and "present tasks for which there is no obvious solution" (by 6 percentage points). In contrast, they are significantly more likely (even though by just 1 percentage point) to report to "have students work in small groups to come up with a joint solution to a problem or task", on average across TALIS countries and economies (see Annex 5.B. for the methodology underlying this estimation).

Teaching approaches and methods

What are the most effective practices for teaching a foreign language?

Teaching practices are an important determinant of learning. For example, across the education systems that participated in PISA 2015, the percentage of science teachers with a major in science was not related to students' proficiency. In contrast, the way science is taught was related not only to science proficiency, but also to how much students value scientific enquiry and to their expectations of working in a science-related occupation (OECD, 2016[37]). The same holds true for (foreign) language learning: the role of the language teacher is central in guiding the students to successful learning (Black and William, 2009[118]; Turner and Purpura, 2015[119]). This section discusses both teaching approaches (broad characterisations of teaching practices reflecting a global understanding of how a language should be taught) and methods (the practical realisation of these approaches in the classroom) (Harmer, 2007[120]).

There is a wide variety of foreign language-teaching practices around the world (Garton, Copland and Burns, 2011[17]). However, research and theory do not define a single "best" way of teaching languages. There is still some controversy, for example, as to whether instruction should be based on a focus-on-forms approach (systematically teaching grammatical features following a structural syllabus), or a focus-on-meaning approach (stimulating learning of linguistic features through communicative activities based on a task-based syllabus (Ellis, 2005[121]). In addition, practices in the classroom do not always follow research ideas or official guidelines (Garton, Copland and Burns, 2011[17]; Renandya et al., 1999[122]).

Traditionally, teaching tended to be teacher-centred, putting the teacher as the expert, instigator and administrator while the student was a passive recipient (Taylor, 2002[123]). In contrast, over the past few decades, education has moved towards a learner-centred approach. In this approach, the teacher's goal is to help students set their own learning goals, manage the content and process of their learning and communicate progressively with peers, also by using ICT tools (Laakkonen, 2011[20]). It is argued that this increases the focus on the needs, circumstances and interests of the learner, improving learning (Lathika, 2016[124]; European Commission, 2017[125]).

The learner-centred approach is particularly relevant to foreign language learning, since active participation gives students opportunities to practice their communicative skills (Sánchez Calvo, 2007[126]). Therefore, the use of a more or less learner-centred approach is a central theme throughout this section. For example,

construct (32) describes communicative language teaching, an approach to student learning that is more learner-centred than traditional teaching. Construct (41) about the use of teaching materials stresses the adoption of practices involving students in the learning process. Group activities, having students use the target language and teacher's talking time (constructs (35), (37) and (38)) also provide ways to measure the extent to which students participate in class. However, it is also important to discuss more traditional teaching practices such as teaching linguistic knowledge (construct (33)).

The system-level questionnaire could collect information about guidelines for language teaching approaches and methods in primary to upper secondary education (as all these levels could affect the cumulative learning of 15-year-old students). Recommending a didactic approach at the national level for teaching a foreign language has a strong impact on classroom practices, even though some studies point to a gap between curriculum policy and classroom practice (Graves and Garton, 2017[127]).

Table 5.9. List of constructs: Teaching approaches and methods

Construct	Policy relevance	Level of measurement
(32) Teaching the four communicative skills	Essential	Student, teacher (teaching practices); school, system (policies and guidelines)
(33) Teaching linguistic knowledge: Grammar, pronunciation, vocabulary	Essential	Student, teacher (teaching practices); school, system (policies and guidelines)
(34) Teaching literature and cultural knowledge	Essential	Student, teacher (teaching practices); school, system (policies and guidelines)
(35) Group and collective learning activities in the classroom	Essential	Student, teacher (teaching practices); school, system (policies and guidelines)
(36) Translanguaging	Relevant	Student, teacher (teaching practices); school, system (policies and guidelines)
(37) Use of the target language during foreign language lessons	Essential	Student, teacher (teaching practices); school, system (policies and guidelines)
(39) Teacher's talking time	Relevant	Student, teacher (teaching practices); school, system (policies and guidelines)
(39) Joint learning of language- and non-language-related content	Essential	Student, teacher (actual practices); school, system (policies and guidelines)
(40) Teaching materials used for target language teaching	Essential	Student, teacher (classroom practices); school, system (policies and guidelines)
(41) Use of teaching materials	Essential	Student, teacher (classroom practices); school, system (policies and guidelines)
(42) Use of the Common European Chapter of Reference for Languages	Essential	Student, teacher (classroom practices); school, system (policies and guidelines)

(32) Teaching the four communicative skills

"Teaching the four communicative skills" refers to the frequency of and amount of classroom time dedicated to teaching students to use the four communicative skills of speaking, writing, reading and listening. This focus on the four skills is a fundamental aspect of the "communicative language teaching" approach. This approach emphasises students' ability to make meaning in different contexts rather than focusing on linguistic knowledge, and aims to develop both productive and receptive skills (Graves and Garton, 2017[127]; Roca Gris, 2015[128]). It also emphasises the integration of the teaching of different skills, for example through activities joining listening and writing or reading and speaking.

Given that proficiency in different skills will be assessed, it could be particularly useful to measure the recurrence and amount of time accorded to the use of each communicative skill (writing, speaking, reading, and listening) by itself or in combination with others. It can be expected that prioritising a certain communicative skill during classroom activities would lead to higher proficiency in that skill.

(33) Teaching linguistic knowledge: Grammar, pronunciation, vocabulary

"Teaching linguistic knowledge: Grammar, pronunciation, vocabulary" refers to the frequency of, and amount of time dedicated in the classroom to teaching structural aspects of the target language ("forms"). Placing emphasis on linguistic knowledge means planning activities and using materials specifically targeted at improving students' pronunciation, and the knowledge of grammar and vocabulary in themselves. This type of exercise entails an emphasis on (for example) grammatical patterns, verbal tenses or vocabulary, sometimes decontextualised from a situation of communication. The type of knowledge gained through these tasks is usually referred to as "metalinguistic knowledge" (Roca Gris, 2015[128]), or knowledge *about* the language. Some authors and practitioners consider certain types of decontextualised activities (e.g. flashcards or grammar analysis) as an important means of improving morphological and syntactic knowledge in a context of limited exposition to foreign languages (Hilton, 2019[129]).

Improving grammar, pronunciation and vocabulary is an important goal of both communicative language teaching and the "teaching linguistic knowledge" approach. However, while in communicative language teaching these goals are pursued through an emphasis on communicative skills and the content of communication (e.g. improving vocabulary through listening to a radio show, without necessarily verifying that the new words have been memorised), in the latter approach more attention is given to the linguistic forms in themselves.

(34) Teaching literature and cultural knowledge

"Teaching literature and cultural knowledge" refers to the frequency of, and amount of time dedicated to, teaching literature and cultural creations (e.g. songs, movies) of target language-speaking communities when teaching the target language. This approach can be combined with different levels of emphasis on linguistic knowledge and the four communicative skills, as both linguistic knowledge and the four communicative skills can be learned while teaching culture and literature. The emphasis on literature and cultural knowledge has been shown to be positively associated with students' foreign language test scores across European countries (European Commission, 2012[31]). In addition, understanding the culture of target language countries can enhance intercultural communication in the target language (Ali, Kazemian and Mahar, 2015[130]).

(35) Group and collective learning activities in the classroom

"Group and collective learning activities in the classroom" refers to the frequency, dedicated time, and type of group activities to learn the target language through interaction with peers in the classroom or for class (e.g. collaborative homework). This variable can measure the size of the groups (pairs or larger) and the level of student autonomy (only limited interaction following a structured template; or more independent work, including researching answers to a question autonomously as a group), and also the use of ICT in these activities. It is important to understand whether the focus of such activities is on language learning or on using language for learning (i.e. to research specific information on a topic on line, filtering it and then reporting).

The interaction and negotiation of meaning that typically occur in group work, when students can autonomously express themselves, are important factors in learning a new language (Brown, 2007[49]; Farrell, 2001[131]). Collaborative activities using a foreign language can also develop mediation capacities, as students have to collaborate to construct meaning, or facilitate interaction with peers, especially in mixed-level group activities (Council of Europe, 2018[77]).

Group and collaborative learning activities are the basis for the communicative language teaching approach (Graves and Garton, 2017[127]; Renandya et al., 1999[122]), and they can be particularly important if opportunities to practice the target language outside of school are scarce.

(36) Translanguaging

"Translanguaging" refers to a teaching approach through which the teachers allow other languages spoken by the students to be used to shape understanding and to be included in foreign language teaching and learning (Alby and Léglise, 2018[132]). Translanguaging has greater relevance in multilingual contexts where minority-language or immigrant students are present in the classroom. In these contexts, students and teachers can better use their "continua of biliteracy", drawing from multiple and dynamic varieties of languages and literacies (Hornberger and Link, 2012[133]).[8] Translanguaging uses all the linguistic resources of the student to maximise understanding (Baker, 2011[134]; Lewis, Jones and Baker, 2012[135]). It is important to understand if translanguaging happens in the classroom, and if it is allowed or encouraged.

The simultaneous use of multiple languages in the classroom is thought to lead to broader and deeper knowledge of language and subjects (Williams, 1996[136]). This approach is also believed to be particularly beneficial for bilingual and multilingual students, as a way to learn and develop language skills using their own resources (Meier and Conteh, 2014[137]). Allowing and encouraging students to speak their other languages is also a way to value and preserve their linguistic and cultural backgrounds.

(37) Use of the target language during foreign language lessons

"Use of the target language during foreign language lessons" indicates the extent to which students practice the target language in the class, but also the extent to which teachers use it to teach (e.g. the frequency of activities involving the use of the target language, or the proportion of time students and teachers speak in the target language as opposed to reading test language).

Group work and conversation in the target language, as well as writing fictitious e-mails or reading newspaper articles in the target language, are ways for students to use the target language during lessons. According to recent research, teachers' use of the target language in the classroom includes three distinct aspects: managing the classroom, understanding and communicating lesson content, and assessing students and giving feedback. For example, giving instructions in the target language as opposed to the language commonly used amongst students is one of the most important requirements of the communicative skills teaching approach (Garton, Copland and Burns, 2011[17]).

(38) Teacher's talking time

"Teacher's talking time" is the amount of time during which the teacher's talks in the classroom while students listen. The teacher can use talking time to explain concepts or manage the classroom, in the target or other language. Teacher's talking time leaves less time to students to practice oral participation and interaction. For example, in some education settings in western China, the large majority of talking time is reserved for teachers and very few students initiate a question in class (Liu, 2016[138]).

(39) Joint learning of language- and non-language-related content

"Joint learning of language- and non-language-related content" applies only to education programmes where the target language is used for instruction in other subjects (see the section *Use of the target language for instruction in other subjects*). It refers to the extent to which teachers purposely integrate the learning of non-language-related content (e.g. mathematics or history) into the learning of the target language. In other words, this construct refers to the "communication" element of Coyle's (1999[139]) chapter for effective learning through a foreign language, i.e. using language to learn while learning to use language. The measurement of this construct may involve asking whether the target language is also a subject of teaching in the lessons of other subjects; and if students are encouraged to ask for the help they need to learn the language, as recommended by experts (Mehisto, Frigols and Marsh, 2008[140]).

(40) Teaching materials used for target language teaching

"Teaching materials used for target language teaching" refers to materials that are used in class and for homework or assignments, and to how often they are used. For analytical purposes, the materials included in this construct should be aligned with those included in "school resources for target language teaching" (construct (8)) to understand how closely use and availability are related. Teaching materials can be divided into three categories:

- textbooks
- ICT tools, such as online platforms, video-sharing websites, foreign language learning software, computer applications and computer-assisted language learning (a tool consisting of online environments where learners can communicate with foreign language speakers, but also online apps, game-based learning, etc. (European Commission, 2014[19]))
- other material such as DVDs, whiteboards, radio and teacher-prepared material.

Some materials may fit some teaching methods better than others. For example, ideally materials for improving students' communicative abilities should be "authentic" (i.e. produced for reasons other than language teaching) and allow meaningful communication in the target language (Brown, 2007[49]; Graves and Garton, 2017[127]; Farrell, 2001[131]). Questions on which teaching materials are used should allow to identify at least some types of materials that are inherently authentic (e.g. movies, news items or blogs in the target language).

ICT tools are becoming an increasingly important part of foreign language lessons, and students need to learn to use these tools for their own future (Al-Mahrooqi and Troudi, 2014[18]; Motteram, 2013[141]; Farr and Murray, 2016[142]). However, no robust evidence has been found across European countries linking student target language proficiency with the use of ICT in the classroom, multimedia language labs, virtual learning environments, the availability of software for language assessment, or use of ICT devices and web content for teaching (European Commission, 2012[31]).

According to the latest data from TALIS, the use of ICT for class work across TALIS countries and economies increased over the past years (OECD, 2019[40]). In 2018, 62% of lower secondary foreign language teachers reported frequently or always letting students use ICT for project or classwork, compared to 57% of other lower secondary teachers (OECD, 2020[21]). However, the data also suggest limited preparation and support available for teachers. Only 60% of teachers reported having received training in the use of ICT for teaching as part of their formal education or training (OECD (2019[40]), Table I.4.13), and this proportion was 2 percentage points lower among foreign language teachers than among other teachers (OECD, 2020[21]).

(41) Use of teaching materials

Interactive use of teaching materials refers to the way in which teaching materials are used in the classroom. This can differ in two important ways:

- There could be more or less interaction amongst students (or between students and teachers) in using the materials proposed by the teacher. For example, teaching materials may be used to give a lecture to students, illustrate linguistic concepts, or present cultural content (e.g. showing a movie to the class). Alternatively, they could be used in more interactive ways, for example if students discuss the material in groups or present to the class content drawn from the teaching materials. For example, teachers can use ICT tools to stimulate students' active participation (e.g. interactions with other students on line) or to support their own teaching (e.g. organising their lesson through a PowerPoint presentation), fulfilling different functions in the language learning process.

- Teachers may adapt their use of textbooks or other materials designed for educational purposes based on their teaching approach and on the class level. For example, they can propose exercises that are based on textbook content but different from those appearing in the textbook. Alternatively, they may use content and exercises as proposed in the materials without adaptation.

Teachers adapting the materials or using them interactively follow a more learner-centred approach, with potentially beneficial effects on student learning (Cruz Rondón and Velasco Vera, 2016[143]).

(42) Use of the Common European Chapter of Reference for Languages

"Use of the Common European Chapter of Reference for Languages" refers to the use of the Council of Europe's standardised chapter (CEFR) (Council of Europe, 2018[77]; Council of Europe, 2001[1]). The CEFR is widely recognised as "the most important reference document in the fields of language learning, teaching, and assessment, both in Europe and beyond" (Barni and Salvati, 2017, p. 417[144]). For example, it is also used to define standard and competencies in the majority of Latin American countries (Cronquist and Fiszbein, 2017[84]) and (with some adaptations) in Japan and a number of ASEAN countries (Foley, 2019[145]). This construct analyses whether teachers use the CEFR for diagnosing students' competence, for teaching (preparing activities, setting goals, etc.) or for evaluation. It is also important to analyse if teachers were trained, during initial or in-service training, to use such chapters. Teachers' training and use of the CEFR for different reasons (teaching, evaluation, etc.) have been found to be positively associated with language proficiency across European countries (European Commission, 2012[31]). If other chapters of reference (e.g. the China Standards of English, Jin et al. (2017[146])) are relevant to countries participating in the assessment, questions on their use could also be included.

Assessment practices

Can national and school target language proficiency assessments improve students' target language learning and proficiency?

The assessment of students' learning and competencies can provide information to students, teachers, governments and other stakeholders of what students have learned and where they stand. This is essential information for (re-) directing the learning process. Students can use this information to change their learning behaviour, teachers to plan their classes, and governments to design education reforms. Assessment can affect teachers' attitudes, teaching content and classroom interactions (Cheng, 2005[147]); when the assessment is "high stakes", for example university entrance exams, it can define the content and performance objectives of education programmes (and it can also affect students' motivation to learn the target language). This section discusses classroom and system-level assessments, and how they can be used to improve student learning. The related question of teachers' assessment literacy is briefly discussed in constructs (26) and (27).

Table 5.10. List of constructs: Assessment practices

Construct	Policy relevance	Level of measurement
(43) Existence of system-level target language assessments	Essential	System
(44) Assessment for learning	Essential	Teacher, student (classroom practices); system (policies and guidelines)

(43) Existence of system-level target language assessments

"Existence of national target language assessments" indicates whether standardised assessments are in place to monitor students' target language proficiency across the education system. It also takes into

consideration assessments in the form of system-level exams that evaluate students' acquisition of curricular content or key competencies.

It is important to determine the scope of these assessments, i.e.:

- whether they are national (e.g. as in China (Zheng and Liying, 2008[148]) or France (Ministère de l'Education Nationale et de la Jeunesse, 2019[149])), or regional (e.g. by province or state, as in Canada (Volante and Ben Jaafar, 2008[150]) or the USA) or some combination, as in Spain (Ministerio de Ciencia, 2018[151]).
- whether all students are required to take them (and whether to progress to a different level, e.g. to graduate or enter higher education) or just a subsample (e.g. as in the National Assessment of Educational Progress in the USA to gather system-wide information for policy making (National Center for Education Statistics, 2018[152])).
- the purpose of the assessment (placement of students into specific programmes or levels, diagnosis of skills attainment, language certification, admission to tertiary education, etc.).
- whether they are based on the curricula taught in schools (e.g. as in China (Zhou and Ito, 2011[153]) or Taiwan (The College Entrance Examination Center, 2019[154])) or intended to function independently of curricula as proficiency tests (e.g. like the International English Language Testing System or the Test of English as a Foreign Language).
- the grades in which students are assessed.
- whether there is a level in the target language students are expected to reach in these tests (whether this level is aligned with the CEFR or not).
- whether they comprehensively address all of the language skills (and how they are weighted) or just some of them (some countries may choose to assess only some skills because of financial constraints; e.g. in 2016 in the *Comunidad de* Madrid in Spain, oral skills were withdrawn from some diagnostic tests in primary education to reduce costs (Enever, 2018[155])).
- whether preparatory activities have been undertaken to prepare to "impact by design", through supporting stakeholders in the testing process, providing them with comprehensive information and monitor and evaluate results (Saville, 2012[156]).

These assessments could affect language proficiency by providing the national and local governments with information used to improve and target their education policies, even if the tests if do not exactly match the learning goals stated in the curriculum.

(44) Assessment for learning

Assessment is "the act of collecting information and making judgements on a language learner's knowledge of a language and ability to use it" (Brindley, 2003[157]). On its own, this is not sufficient for improving learners' proficiency. The information collected must also provide information on how to improve student performance, and the signals it conveys must be acted upon. The UK-based Assessment Reform Group defined "assessment for learning" as "the process of seeking and interpreting evidence for use by learners and their teachers to decide where the learners are in their learning, where they need to go and how best to get there" (Wiliam, 2011, p. 10[158]).[9] The interpretation of student performance evidence and its use to help learners progress are also the fundamental elements in the definition of "learning oriented assessment" (Jones and Saville, 2016, p. 2[159]).

Based on this definition, understanding whether assessments are used to support learning requires information about the assessment itself and on how it is used (Wiliam, 2011[158]). It is necessary to know

how, how often, by whom and with what consequences proficiency in the target language is assessed at school. Assessments can be conducted by teachers, peers and the individual students themselves, with different modalities (e.g. vocabulary or grammar tests, project work, essay writing, dialogues, self-assessment and records (Faustino, Kostina and Vergara, 2013[160]; Brown and Hudson, 1998[161]; Council of Europe, 2001[1])). Teachers should provide constructive and encouraging feedback (Jones and Saville, 2016[159]), and the information provided by the assessment should feed back into their teaching and course planning (Wiliam, 2011[158]; Faustino, Kostina and Vergara, 2013[160]). This implies that teachers understand how to use assessment results (Fulcher, 2012[162]; Hasselgreen, 2005[163]; Coombe, Troudi and Al-Hamly, 2012[107]), which can be part of their initial or in-service training (constructs (26) and (27)).

Notes

1. Countries were asked to rate each variable as “essential”, “relevant” or “not directly relevant”. These ratings were described to countries as follows. “Essential” means that excluding this variable from the [context] questionnaires seriously undermines the interpretability and/or usefulness of the data produced through the PISA Foreign Language Assessment. For example, you may rate a context variable “essential” because you think that without this information it will not be possible to compare the proficiency scores across participating countries. Or you may rate a variable on teaching or school practices as “essential” because your country is implementing a major reform involving this variable, and this is one of the key motives for you to participate in the PISA Foreign Language Assessment. “Relevant” means that this variable is of interest for your country. This could be because it enhances the interpretability of the data or because it can inform the national policy debate. For example, a variable is relevant if it has been brought up in national policy discussions by the media, unions, experts or public authorities. In addition, a context variable can be relevant if it helps comparing the proficiency scores across countries. “Not directly relevant” means that this variable can be excluded from the [context] questionnaire without important consequences on the comparability of the proficiency scores or on their usefulness to policy analysis. For example, a variable is not directly relevant if you do not think it relates to target language learning or proficiency, and you cannot think of any insightful analysis that could be conducted with it.

2. This rule for the determination of construct policy relevance has three exceptions. Constructs (22) and (23) (rated as “relevant” in this chapter) were part of a list of constructs on “intercultural and interdisciplinary capabilities and their relation to foreign language proficiency” that received low ratings in the survey. The countries provided substantial and constructive input to revise the construct at the Paris workshop, and concluded that the constructs would be relevant or essential after the revision. Construct (38) (rated as “relevant” in this chapter) did not receive a rating through the survey because it was not included in the first version of the chapter. Workshops participants did not consider this construct necessary, considering it as implied by other constructs in the chapter. However, it has been included in the current draft following the recommendation of several reviewers.

3. The survey was returned by Chile, Colombia, the Flemish Community of Belgium, France, Italy, Luxembourg, the Netherlands, Portugal, Spain, Switzerland, Ukraine and the United Arab Emirates. Representatives from Colombia, the Flemish Community of Belgium, France, Germany, Italy, Luxembourg, the Netherlands, Portugal, Spain, Switzerland and the United Arab Emirates participated in the workshop in Paris. The PISA Governing Board, which comprises all countries that participate in PISA, welcomed the chapter and the ratings at its meeting in September 2019.

4. This list of tools is presented as an illustrative example. It does not aim to be comprehensive, nor to include the most widely used or researched tools. There is an enormous number of ICT tools already available to teachers and learners, and their design and names changes across users and evolve through time.

5. The policy relevance of construct (7) “Target language remedial lessons at school” was changed from “Relevant” to “Essential” based on a second construct rating provided by 15 PISA countries and economies in September 2020. With the exception of Construct (7), these ratings confirmed the relevance agreed upon in 2019 (see section “Determination of the policy relevance of the constructs”).

6. In the FLA Framework, “task” is used to refer to any purposeful language activity that language users and learners engage with; this includes, in the context of the foreign language test, any assessment items or exercises (e.g. a set of questions about a text), which are also referred to as (assessment) tasks.

7. The relationship between foreign language learning and intercultural understanding is distinct from the relationship between foreign language learning and interest in a specific culture (see the notion of “foreignness” of English as a lingua franca (Seidlhofer, 2011[91])). The latter relationship arises if interest in a certain culture motivates students to learn an associated language (or if learning the language deepens their interest in a related culture).

8. This is in contrast to situations in which the dominant language within a community is used systematically in foreign language teaching (e.g. because of low teacher proficiency in the target language), potentially hindering students’ development of communicative skills, such as speaking and listening.

9. As a concept, “assessment for learning” is related to (and sometimes used interchangeably with) formative assessment. Assessment is formative when the information on student performance is used to adapt teaching to meet students’ needs (Black and Wiliam, 1998, in Wiliam (2011, p. 9[158])). Formative assessment is usually contrasted with summative assessment or “assessment of learning”, which “aims to summarise learning that has taken place, in order to record, mark or certify achievements” (OECD, 2013, p. 140[174]). This chapter adopts a broad definition of “assessment for learning” which avoids the distinction between formative and summative assessment. This distinction is not clear, practically and conceptually. There are strong complementarities between the two types of assessment, and formative assessment can be used for summative purposes (and vice versa) (Jones and Saville, 2016[159]).

References

Ainley, J. and R. Carstens (2018), "Teaching and Learning International Survey (TALIS) 2018 Conceptual Framework", *OECD Education Working Papers*, No. 187, OECD Publishing, Paris, https://dx.doi.org/10.1787/799337c2-en. [5]

Alby, S. and I. Léglise (2018), "Multilingualism and translanguaging as a resource for teaching and learning in French Guiana", in Van Avermaet, P. et al. (eds.), *The multilingual edge of education*, Palgrave Macmillan, https://hal.archives-ouvertes.fr/hal-01692137/document. [132]

Ali, S., B. Kazemian and I. Mahar (2015), "The importance of culture in second and foreign language learning", *Dinamika Ilmu*, Vol. 15/1. [130]

Al-Mahrooqi, R. and S. Troudi (2014), *Using Technology in Foreign Language Teaching*, Cambridge Scholars Publishing, https://www.cambridgescholars.com/download/sample/61911. [18]

Amirabadi, S. and S. Razmjoo (2017), "Social, Cultural and Linguistic Factors Affecting EFL Learners' Language Proficiency: A Quantitative Analysis", *International Journal of Foreign Language Teaching & Research*, Vol. 5/17, pp. 121-135. [51]

Aoumeur, H. (2017), "The Impact of Class Size on Teaching and Learning English as a Foreign Language: The Case of the Department of English at Abdelhamid Ibn Badis University", *Arab World English Journal (AWEJ)*, Vol. 8/2, p. 2017, http://dx.doi.org/10.24093/awej/vol8no2.25. [39]

Baker, C. (2011), *Foundations of bilingual education and bilingualism*, Multilingual Matters, Clevedon, UK. [134]

Barac, R. and E. Bialystok (2010), "Cognitive development of bilingual children", *Language Teaching*, Vol. 44/1, pp. 36-54, http://dx.doi.org/10.1017/s0261444810000339. [47]

Barni, M. and L. Salvati (2017), "The Common European Framework of Reference (CEFR)", in Shohamy, E., I. Or and S. May (eds.), *Language Testing and Assessment. Encyclopedia of Language and Education (3rd ed.)*, Springer, Cham, https://doi.org/10.1007/978-3-319-02261-1_29. [144]

Bartram, B. (2018), *Attitudes to Modern Foreign Language Learning – Insights from Comparative Education*, Bloomsbury, London, https://www.bloomsbury.com/uk/attitudes-to-modern-foreign-language-learning-9781441192882/. [59]

Berns, M., K. de Bot and U. Hasebrink (eds.) (2007), *In the Presence of English: Media and European Youth*, Springer US, Boston, MA, http://dx.doi.org/10.1007/978-0-387-36894-8. [78]

Black, P. and D. William (2009), "Developing the theory of formative assessment", *Educational Assessment, Evaluation and Accountability*, Vol. 21/1, pp. 5-31. [118]

Bonnet, G. (ed.) (2002), *Assessment of Pupils' Skills in English in Eight European Countries*, Ministère de l'Éducation Nationale, Paris, http://www.educacionyfp.gob.es/inee/dam/jcr:d426c134-5c58-44f0-8641-3b5e4354ed37/habilidadesingles2002-1.pdf. [56]

Brindley, G. (2003), "Classroom-based assessment", in Nunan, D. (ed.), *Practical English Language Teaching*, McGraw Hill, New York. [157]

Brown, A. (2009), "Students' and Teachers' Perceptions of Effective Foreign Language Teaching: A Comparison of Ideals", *The Modern Language Journal*, Vol. 93/1, pp. 46-60. [50]

Brown, H. (2007), *Teaching by principles*, Pearson Education, NY. [49]

Brown, J. and T. Hudson (1998), "The Alternatives in Language Assessment", *TESOL Quarterly*, Vol. 32/4, p. 653, http://dx.doi.org/10.2307/3587999. [161]

Bruton, A. (2013), "CLIL: Some of the Reasons Why... And Why Not", *System*, Vol. 41/3, pp. 587-597, http://dx.doi.org/10.1016/j.system.2013.07.001. [26]

Bruton, A. (2011), *Are the Differences between CLIL and Non-CLIL Groups in Andalusia Due to CLIL? A Reply to Lorenzo, Casal and Moore (2010)*, http://dx.doi.org/10.1093/applin/amr007. [105]

Burstall, C. (1975), "Factors Affecting Foreign-Language Learning: A Consideration of Some Recent Research Findings", *Language Teaching & Linguistics*, Vol. 8/1, pp. 5-25, http://dx.doi.org/10.1017/S0261444800002585. [92]

Butler, Y. (2004), "What level of English proficiency do elementary school teachers need to attain to teach EFL? Case studies from Korea, Taiwan, and Japan", *TESOL Quarterly*, Vol. 38/2, pp. 245-278. [115]

Butler, Y., Y. Someya and E. Fukuhara (2014), "Online games for young learners' foreign language learning", *ELT Journal*, Vol. 68/3, pp. 265-275, http://dx.doi.org/10.1093/elt/ccu008. [16]

Cambridge Assessment (2017), *Bilingual learners and bilingual education*, Cambridge Assessment International Education, Cambridge, http://www.cambridgeenglish.org. [24]

Carroll, J. (1975), *The Teaching of French in Eight Countries*, John Wiley & Sons, New York, http://dx.doi.org/prova123. [12]

Carroll, J. (1963), "A Model of School Learning", *Teachers College Record*, Vol. 64, pp. 723-733. [45]

Cenoz, J. (2003), "The Additive Effect of Bilingualism on Third Language Acquisition: A Review", *International Journal of Bilingualism*, Vol. 7/1, pp. 71-87, https://doi.org/10.1177/13670069030070010501. [48]

Chang, K. (2012), "Language Education in Japan and Korea: Policies, Practices and Challenges", in Della Chiesa, B., J. Scott and C. Hinton (eds.), *Languages in a Global World: Learning for Better Cultural Understanding*, OECD Publishing, Paris, http://dx.doi.org/10.1787/9789264123557. [102]

Cheng, L. (2005), *Changing language teaching through language testing: a washback study*, Cambridge University Press, Cambridge. [147]

Coombe, C., S. Troudi and M. Al-Hamly (2012), "Foreign and Second Language teacher assessment literacy: Issues, challenges and recommendations", in Sullivan, B. and S. Stoynoff (eds.), *The Cambridge Guide to Second Language Assessment*, Cambridge University Press, NY, USA. [107]

Cooper, H., J. Robinson and E. Patall (2006), "Does Homework Improve Academic Achievement? A Synthesis of Research, 1987–2003", *Review of Educational Research*, Vol. 76/1, pp. 1-62, http://dx.doi.org/10.3102/00346543076001001. [81]

Cordingley, P. et al. (2015), *Developing Great Teaching: Lessons from the international reviews into effective professional development*, Teacher Development Trust. [108]

Council of Europe (2018), *Common European Framework of Reference for Languages: Learning, Teaching, Assessment - Companion Volume with New Descriptors*, Council of Europe, Strasbourg, http://www.coe.int/lang-cefr. [77]

Council of Europe (2008), *White Paper on Intercultural Dialogue "Living Together As Equals in Dignity"*, Council of Europe, Strasbourg, http://www.coe.int/dialogue. [88]

Council of Europe (2001), *Common European Framework of Reference for Languages: Learning, Teaching, Assessment*, Cambridge University Press, Cambridge. [1]

Council of the European Union (2019), "Council Recommendation of 22 May 2019 on a Comprehensive Approach to the Teaching and Learning of Languages", *Official Journal of the European Union*, Vol. C 189, pp. 15-22, http://dx.doi.org/prova123. [43]

Coyle, D. (1999), "Theory and planning for effective classrooms: supporting students in content and language integrated learning contexts", in Masih, J. (ed.), *Learning through a foreign language*, CILT, the National Centre for Languages. [139]

Cronquist, K. and A. Fiszbein (2017), *El Aprendizaje del Inglés en América Latina*, The Inter-American Dialogue, Washington, D.C., https://www.thedialogue.org/wp-content/uploads/2017/09/El-aprendizaje-del-ingl%C3%A9s-en-Am%C3%A9rica-Latina-1.pdf. [84]

Cruz Rondón, E. and L. Velasco Vera (2016), "Understanding the Role of Teaching Materials in a Beginners' Level English as a Foreign Language Course: A Case Study", *PROFILE*, Vol. 18/2, pp. 1657-0790, http://www.scielo.org.co/pdf/prf/v18n2/v18n2a08.pdf. [143]

Cuenat, E., L. Bleichenbacher and R. Frehner (2016), "Plurilingual and Intercultural Education in Teacher Training Mobility: Experiencing and Implementing PluriMobil.", in *Teacher Education in the 21st Century: A Focus on Convergence*, Cambridge Scholar Publishing, Newcastle upon Tyne. [111]

Cumming, A. (ed.) (2012), *Adolescent Literacies in a Multicultural Context*, Routledge, New York, http://dx.doi.org/10.4324/9780203120033. [57]

Cumming, A. (1996), "IEA's Studies of Language Education: Their Scope and Contributions", *Assessment in Education*, Vol. 3, pp. 179-192, https://doi.org/10.1080/0969594960030205. [13]

Cummins, J. (2005), "A Proposal for Action: Strategies for Recognizing Heritage-language Competence as a Learning Resource within the Mainstream Classroom", *The Modern Language Journal*, Vol. 98, pp. 585-592, https://www.jstor.org/stable/3588628. [100]

Dafouz, E. and A. Hibler (2013), "'Zip Your Lips' or 'Keep Quiet': Main Teachers' and Language Assistants' Classroom Discourse in CLIL Settings", *Modern Language Journal*, Vol. 97/3, pp. 655-669, http://dx.doi.org/10.1111/j.1540-4781.2013.12026.x. [106]

Dallinger, S. et al. (2016), "The effect of content and language integrated learning on students' English and history competences - Killing two birds with one stone?", *Learning and Instruction*, Vol. 41, pp. 23-31, http://dx.doi.org/10.1016/j.learninstruc.2015.09.003. [171]

Della Chiesa, B. (ed.) (2012), *Learning Languages in a Globalising World*, Languages in a Global World: Learning for Better Cultural Understanding, OECD Publishing, Paris, https://dx.doi.org/10.1787/9789264123557-en. [89]

Denies, K. and R. Janssen (2016), "Country and Gender Differences in the Functioning of CEFR-Based Can-Do Statements as a Tool for Self-Assessing English Proficiency", *Language Assessment Quarterly*, Vol. 13/3, pp. 251-276, http://dx.doi.org/10.1080/15434303.2016.1212055. [79]

Derakhshan, A. and E. Karimi (2015), "The Interference of First Language and Second Language Acquisition", *Theory and Practice in Language Studies*, Vol. 5/10, p. 2112, http://dx.doi.org/10.17507/tpls.0510.19. [52]

Donato, R. (1994), "Collective scaffolding in second language learning", in Lantolf, J. and G. Appel (eds.), *Vygotskian Approaches to Second Language Research*, Ablex, Norwood, NJ. [63]

Dörnyei, Z. (2010), "The Relationship Between Language Aptitude and Language Learning Motivation: Individual Differences from a Dynamic Systems Perspective", in Macaro, E. (ed.), *Continuum Companion to Second Language Acquisition*, MPG Books Group, Bodmin. [66]

Dörnyei, Z. (2005), *The Psychology of the Language Learner: Individual Differences in Second Language Acquisition*, Lawrence Erlbaum Associates, Mahwah, http://dx.doi.org/prova123. [67]

Dörnyei, Z. and L. Chan (2013), "Motivation and Vision: An Analysis of Future L2 Self Images, Sensory Styles, and Imagery Capacity Across Two Target Languages", *Language Learning*, Vol. 63/3, pp. 437-462, http://dx.doi.org/10.1111/lang.12005. [75]

Dörnyei, Z. and P. Skehan (2003), "Individual Differences in Second Language Learning", in Doughty, C. and M. Long (eds.), *The Handbook of Second Language Acquisition*, Blackwell Publishing Ltd, Oxford, UK, http://dx.doi.org/10.1002/9780470756492.ch18. [71]

Ellis, R. (2005), "Principles of instructed language learning", *The Asian EFL Journal*, Vol. 7/3, pp. 9-24, http://asian-efl-journal.com/sept_05_re.pdf. [121]

Enever, J. (2018), *Policy and politics in global primary English*, Oxford University Press, Oxford. [155]

Enever, J. (ed.) (2011), *ELLiE Early Language Learning in Europe*, British Council. [10]

European Commission (2017), *Rethinking Language Education and Linguistic Diversity in Schools*, Publications Office of the European Union, Luxembourg, http://dx.doi.org/10.2766/584023. [125]

European Commission (2015), *Language and Learning in Multilingual Classrooms*, Publications Office of the European Union, Luxembourg, http://dx.doi.org/10.2766/766802. [98]

European Commission (2014), *Improving the effectiveness of language learning: CLIL and computer assisted language learning*, Education and Training, Bruxelles, https://www.ecml.at/Portals/1/resources/Articles%20and%20publications%20on%20the%20ECML/CLIL%20and%20CALL%20report_July.2014.pdf?ver=2017-07-11-151504-977. [19]

European Commission (2012), *First European Survey on Language Competences: Final Report*, Publications Office of the European Union, Bruxelles, https://www.researchgate.net/publication/262877352. [31]

European Commission (2012), *First European Survey on Language Competences: Technical Report*, European Commission, Bruxelles, http://ec.europa.eu/assets/eac/languages/policy/strategic-framework/documents/language-survey-technical-report_en.pdf. [6]

European Commission (2006), *Content and Language Integrated Learning (CLIL) at school in Europe*, Eurydice, Brussels, https://publications.europa.eu/en/publication-detail/-/publication/756ebdaa-f694-44e4-8409-21eef02c9b9b. [168]

European Commission (1995), *White paper on education and training. Teaching and learning: towards the learning society*, European Commission, Brussels, http://europa.eu/documents/comm/white_papers/pdf/com95_590_en.pdf. [167]

European Commission/EACEA/Eurydice (2017), *Key Data on Teaching Languages at School in Europe - 2017 Edition*, Publications Office of the European Union, Luxembourg, http://dx.doi.org/10.2797/839825. [11]

Fan, H. et al. (2017), "Homework and students' achievement in math and science: A 30-year meta-analysis, 1986–2015", *Educational Research Review*, Vol. 20, pp. 35-54, http://dx.doi.org/10.1016/j.edurev.2016.11.003. [82]

Farrell, T. (2001), "Paradigm shift: Understanding and implementing change in second language education", *TESL-EJ*, Vol. 5/1, http://www.esletc.com/2007/08/19/jacobs-george-m-thomas-s-c-farrell-2001-paradigm-shift-understanding-and-implementing-change-in-second-language-education-tesl-ej-51/. [131]

Farr, F. and L. Murray (2016), *The Routledge Handbook of Language Learning and Technology*, Routledge. [142]

Faustino, C., I. Kostina and O. Vergara (2013), "Assessment Practices in the English and French Component of a Foreign Languages Teacher Education Program", *Lenguaje*, Vol. 41/2, pp. 353-382, http://www.scielo.org.co/scielo.php?script=sci_arttext&pid=S0120-34792013000200004. [160]

Flemish Ministry of Education and Training (2019), *Frans Lager Onderwijs*, https://www.kwalificatiesencurriculum.be/frans-lager-onderwijs (accessed on 23 July 2019). [69]

Foley, J. (2019), "Adapting CEFR for English Language Education in ASEAN, Japan and China", *The New English Teacher* 2, pp. 101-117, http://www.assumptionjournal.au.edu/index.php/newEnglishTeacher/article/view/3879/2371. [145]

Freeman, D. et al. (2015), "English-for-Teaching: Rethinking teacher language proficiency for the classroom", *English Language Teaching Journal*, Vol. 69/2, pp. 129-139, http://dx.doi.org/10.1093/elt/ccu074. [117]

Fulcher, G. (2012), "Assessment Literacy for the Language Classroom", *Language Assessment Quarterly*, Vol. 9/2, pp. 113-132, http://dx.doi.org/10.1080/15434303.2011.642041. [162]

Gaonac'h, D. and D. Macaire (2019), *Le Langues a l'École dès les Plus Jeune Âge*, Conseil National d'Évaluation du Système Scolaire (CNESCO), Paris. [36]

Gardner, R. (1985), *Social Psychology and Second Language Learning: The Role of Attitude and Motivation*, Hodder Arnold, Baltimore. [70]

Garrett, N. (2009), "Computer-Assisted Language Learning Trends and Issues Revisited: Integrating Innovation", *The Modern Language Journal*, Vol. 93, pp. 719-740, http://dx.doi.org/10.1111/j.1540-4781.2009.00969.x. [15]

Garton, S., F. Copland and A. Burns (2011), *Investigating Global Practices in Teaching English to Young Learners*, British Council, London, https://www.teachingenglish.org.uk/sites/teacheng/files/B094%20FINAL%20Aston%20University%20A4%20report_2column_V3.pdf. [17]

Graddol, D. (2006), *English Next*, British Council, London, https://englishagenda.britishcouncil.org/sites/default/files/attachments/books-english-next.pdf. [99]

Grant, J. (2012), "Music as an Underutilised and Underappreciated Tool for Language Learning", in Della Chiesa, B., J. Scott and C. Hinton (eds.), *Languages in a Global World: Learning for Better Cultural Understanding*, OECD Publishing, Paris, http://dx.doi.org/10.1787/9789264123557-en. [62]

Graves, K. and S. Garton (2017), "An analysis of three curriculum approaches to teaching English in public-sector schools", *Language Teaching*, Vol. 50/4, pp. 441-482, http://dx.doi.org/10.1017/S0261444817000155. [127]

Hanesová, D. (2015), "History of CLIL", in Pokrivčáková, S. (ed.), *CLIL in Foreign Language Education: e-textbook for foreign language teachers*, Constantine the Philosopher University, Nitra. [164]

Hanushek, E. and S. Rivkin (2004), "How to Improve the Supply of High-Quality Teachers", *Brookings Papers on Education Policy*, Brookings, Washington. [103]

Harmer, J. (2007), *The practice of English Language Teaching*, Longman. [120]

Hartshorne, J., J. Tenenbaum and S. Pinker (2018), "A critical period for second language acquisition: Evidence from 2/3 million English speakers", *Cognition*, Vol. 177, pp. 263-277, http://dx.doi.org/10.1016/J.COGNITION.2018.04.007. [30]

Hasselgreen, A. (2005), "Assessing the language of young learners", *Language Testing*, Vol. 223/3, pp. 337-354, http://dx.doi.org/10.1191/0265532205lt312oa. [163]

Heath, L. (2015), "Triangulation: Methodology", in *International Encyclopedia of the Social & Behavioral Sciences*, Elsevier, http://dx.doi.org/10.1016/b978-0-08-097086-8.44059-6. [14]

Heugh, K. (2018), "Conclusion: Multilingualism, Diversity and Equitable Learning", in Van Avermaet, P. et al. (eds.), *The multilingual edge of education*, Palgrave Macmillan, London, http://dx.doi.org/10.1057/978-1-137-54856-6. [97]

Hilton, H. (2019), *Sciences Cognitives et Apprentissage des Langues*, Conseil National d'Évaluation du Système Scolaire (CNESCO), Paris, http://www.cnesco.fr/wp-content/uploads/2019/04/190409_Hilton-1.pdf. [129]

Hoover-Dempsey, K. et al. (2001), "Parental Involvement in Homework", *Educational Psychologist*, Vol. 36/3, pp. 195-209, http://dx.doi.org/10.1207/s15326985ep3603_5. [58]

Hornberger, N. and H. Link (2012), "Translanguaging and transnational literacies in multilingual classrooms: a biliteracy lens", *International Journal of Bilingual Education and Bilingualism*, Vol. 15/3, pp. 261-278, http://dx.doi.org/10.1080/13670050.2012.658016. [133]

Horwitz, E. (1990), "Attending to the affective domain in the foreign language classroom", in Magnan, S. (ed.), *Shifting the instructional focus to the learner*, Northeast Conference on the Teaching of Foreign Languages, Middlebury, VT. [112]

IEA (1993), *The IEA Language Education Study Proposal*, 34th IEA General Assembly, El Escorial. [9]

Isphording, I. and S. Otten (2014), "Linguistic Barriers in the Destination Language Acquisition of Immigrants", *Journal of Economic Behavior and Organization*, No. 8090, http://anon-ftp.iza.org/dp8090.pdf. [54]

Jenkins, J. (2017), "Not English but English-within-Multilingualism", in *New Directions for Research in Foreign Language Education*, Routledge, http://dx.doi.org/10.4324/9781315561561-5. [86]

Jin, Y. et al. (2017), "Developing the China Standards of English: challenges at macropolitical and micropolitical levels", *Language Testing in Asia*, Vol. 7/1, http://dx.doi.org/10.1186/s40468-017-0032-5. [146]

Jones, N. and N. Saville (2016), *Learning Oriented Assessment*, Cambridge University Press, Cambridge. [159]

Kelly, M. (2019), "Language Awareness in a Comprehensive Approach to Language Teaching and Learning", *Journal of E-Learning and Knowledge Society*, Vol. 15/1, pp. 37-54, http://dx.doi.org/10.20368/1971-8829/1616. [42]

Kern, R. (1995), "Students and teachers' beliefs about language learning", *Foreign Language Annals*, Vol. 28, pp. 71-92. [113]

Kim, Y. and T. Kim (2011), "The Effect of Korean Secondary School Students' Perceptual Learning Styles and Ideal L2 Self on Motivated L2 Behavior and English Proficiency", *Korean Journal of English Language and Linguistics*, Vol. 11/1, pp. 21-42, http://dx.doi.org/10.15738/kjell.11.1.201103.21. [60]

King, L. (2017), *The Impact of Multilingualism on Global Education and Language Learning*, Cambridge English Assessment, Cambridge, https://assets.cambridgeenglish.org/research/perspectives-multilingualism.pdf. [87]

Kuchah, K. (2009), "Early bilingualism in Cameroon: Where politics and education meet", in Enever, J., J. Moon and U. Raman (eds.), *Young Learner English Language Policy and Implementation: International Perspectives*, Garnet Education, Reading. [116]

Laakkonen, I. (2011), "Personal learning environments in higher education language courses: an informal and learner-centred approach", in Thouësny S and Bradley L (eds.), *Second language teaching and learning with technology: views of emergent researchers*, Research-publishing.net, Dublin. [20]

Lantolf, J. (2000), *Sociocultural Theory and Second Language Learning*, Oxford University Press, Oxford. [64]

Larson-Hall, J. (2008), "Weighing the benefits of studying a foreign language at a younger starting age in a minimal input situation", *Second Language Research*, Vol. 24/1, pp. 35-63, http://dx.doi.org/10.1177/0267658307082981. [38]

Lasagabaster, D. (2008), "Foreign Language Competence in Content and Language Integrated Courses", *The Open Applied Linguistics Journal* 1, pp. 31-42, https://www.unifg.it/sites/default/files/allegatiparagrafo/20-01-2014/lasagabaster_foreign_language_competence_and_language_integrated_courses.pdf. [22]

Lathika, K. (2016), "Student Centred Learning", *International Journal of Current Research and Modern Education*, Vol. 1/1, pp. 677-680. [124]

Lee, I., S. Yu and Y. Liu (2018), "Hong Kong Secondary Students' Motivation in EFL Writing: A Survey Study", *TESOL Quarterly*, Vol. 52/1, pp. 176-187, http://dx.doi.org/10.1002/tesq.364. [72]

Leigh, A. (2010), "Estimating Teacher Effectiveness from Two-year Changes in Students' Test Scores", *Economics of Education Review*, Vol. 29/3, pp. 480-488, http://dx.doi.org/10.1016/j.econedurev.2009.10.010. [104]

Leppänen, S. et al. (2011), *National Survey on the English Language in Finland: Uses, Meanings and Attitudes*, University of Jyväskylä, Jyväskylä, http://www.helsinki.fi/varieng/journal/volumes/05. [65]

Lewis, E. and C. Massad (1975), *The Teaching of English as a Foreign Language in Ten Countries (International studies in evaluation)*, John Wiley & Sons Inc. [7]

Lewis, G., B. Jones and C. Baker (2012), "Translanguaging: developing its conceptualisation and contextualisation", *Educational Research and Evaluation*, Vol. 18/7, pp. 655-670. [135]

Lightbown, P. and N. Spada (2013), *How Languages are Learned*, Oxford University Press, Oxford. [44]

Lindgren, E. and C. Muñoz (2013), "The Influence of Exposure, Parents, and Linguistic Distance on Young European Learners' Foreign Language Comprehension", *International Journal of Multilingualism*, Vol. 10/1, pp. 105-129, http://dx.doi.org/10.1080/14790718.2012.679275. [53]

Little, D. (2005), "The Common European Framework and the European Language Portfolio: involving learners and their judgements in the assessment process", *Language Testing*, Vol. 22/3, pp. 321-336, http://dx.doi.org/10.1191/0265532205lt311oa. [80]

Liu, L. (2016), "Research on Teacher-Student Interaction in Senior Schools English Teaching in Western China", *US-China Foreign Language*, Vol. 14/5, http://dx.doi.org/10.17265/1539-8080/2016.05.003. [138]

Marsh, H., K. Hau and C. Kong (2000), "Late Immersion and Language of Instruction in Hong Kong High Schools: Achievement Growth in Language and Nonlanguage Subjects", *Harvard Educational Review*, Vol. 70/3, pp. 302-345, http://dx.doi.org/10.17763/haer.70.3.gm047588386655k5. [29]

Marsh, H., K. Hau and C. Kong (2000), "Late Immersion and Language of Instruction in Hong Kong High Schools: Achievement Growth in Language and Nonlanguage Subjects", *Harvard Educational Review*, Vol. 70/3, pp. 302-347, http://dx.doi.org/10.17763/haer.70.3.gm047588386655k5. [173]

Mehisto, P. (2012), *Excellence in Bilingual Education: A Guide for School Principals*, Cambridge University Press, Cambridge. [25]

Mehisto, P., M. Frigols and D. Marsh (2008), *Uncovering CLIL*, Macmillan English. [140]

Mehl, M. (2003), *Private Academies of Chinese Learning in Meiji Japan*, Nordic Institute of Asian Studies, Copenhagen. [165]

Meier, G. and J. Conteh (2014), "Conclusion", in Meier, G. and J. Conteh (eds.), *The multilingual turn in language education: Opportunities and challenges*, Multilingual Matters, Bristol, UK. [137]

Melzi, G. and A. Schick (2012), "Motivation and Second Language Acquisition", in Della Chiesa, B., J. Scott and C. Hinton (eds.), *Languages in a Global World: Learning for Better Cultural Understanding*, OECD Publishing, Paris, http://dx.doi.org/10.1787/9789264123557-en. [74]

Mercer, S. (2011), *Towards an Understanding of Language Learner Self-Concept*, Springer, Berlin. [68]

Ministère de l'Education Nationale et de la Jeunesse (2019), *Voies de formation et diplômes. Le baccalauréat général.*, https://www.education.gouv.fr/cid145/le-baccalaureat-general.html (accessed on 26 July 2019). [149]

Ministerio de Ciencia, I. (2018), *Pruebas de Acceso a la Universidad*, http://www.educacionyfp.gob.es/servicios-al-ciudadano/estadisticas/universitaria/estadisticas/estadistica-de-las-pruebas-de-acceso-a-la-universidad0/ano-2017.html (accessed on 26 July 2019). [151]

Motteram, G. (2013), *Innovations in learning technologies for English language teaching*, British Council, London, http://www.britishcouncil.org. [141]

Muñoz, C. (2006), *Age and the rate of foreign language learning*, Multilingual Matters, Toronto, https://books.google.fr/books/about/Age_and_the_Rate_of_Foreign_Language_Lea.html?id=1C_-zfVkmOkC&redir_esc=y. [33]

Murphy, V. (2014), *Second Language Learning in the Early School Years: Trends and Context*, Oxford University Press, Ofxord. [35]

National Center for Education Statistics (2018), *National Assessment of Educational Progress*, https://nces.ed.gov/nationsreportcard/about/ (accessed on 26 July 2019). [152]

Nieto Moreno de Diezmas, E. (2016), "The impact of CLIL on the acquisition of L2 competences and skills in primary education", *International Journal of English Studies*, Vol. 16/2, p. 81, http://dx.doi.org/10.6018/ijes/2016/2/239611. [170]

Nikolov, M. (ed.) (2016), *Learning EFL from Year 1 or Year 3? A Comparative Study on Children's EFL Listening and Reading Comprehension at the End of Primary Education*, Springer International Publishing, Cham, http://dx.doi.org/10.1007/978-3-319-22422-0. [32]

Nikula, T. (2016), "CLIL: A European Approach to Bilingual Education", in *Second and Foreign Language Education*, Springer International Publishing, Cham, http://dx.doi.org/10.1007/978-3-319-02323-6_10-1. [23]

Nikula, T. (2016), "CLIL: A European Approach to Bilingual Education", in *Second and Foreign Language Education*, Springer International Publishing, Cham, http://dx.doi.org/10.1007/978-3-319-02323-6_10-1. [28]

Nikula, T., C. Dalton-Puffer and A. García (2013), "CLIL classroom discourse: Research from Europe", *Journal of Immersion and Content-Based Language Education*, Vol. 1/1, pp. 70-100, http://dx.doi.org/10.1075/jicb.1.1.04nik. [169]

Noels, K. et al. (2003), "Why Are You Learning a Second Language? Motivational Orientations and Self-Determination Theory", *Language Learning*, Vol. 53, pp. 33-64, http://dx.doi.org/10.1111/1467-9922.53223. [73]

Nor, F., M. Aziz and K. Jusoff (2011), "Should English for Teaching Mathematics and Science (ETeMS) in Malaysia Be Abolished?", *World Applied Sciences Journal* 12, pp. 36-40. [172]

Norton, B. and K. Toohey (2001), "Changing Perspectives on Good Language Learners", *TESOL Quarterly*, Vol. 35/2, pp. 307-322, http://dx.doi.org/10.2307/3587650. [93]

Nunan, D. (2003), "The impact of English as a global language on educational policies and practices in the Asia-Pacific region", *TESOL Quarterly*, Vol. 37/4, pp. 589-613. [83]

OECD (2020), "Foreign language teachers as ambassadors of multilingualism and international exchange: Evidence from TALIS 2018", *Teaching in Focus*, No. 33, OECD Publishing, Paris, https://dx.doi.org/10.1787/808f11a4-en. [21]

OECD (2019), *PISA 2018 Assessment and Analytical Framework*, PISA, OECD Publishing, Paris, https://dx.doi.org/10.1787/b25efab8-en. [94]

OECD (2019), "PISA 2018 Global Competence Framework", in *PISA 2018 Assessment and Analytical Framework*, OECD Publishing, Paris, https://dx.doi.org/10.1787/043fc3b0-en. [85]

OECD (2019), *PISA 2021 Context Questionnaires Framework (Field Trial Version)*, OECD, Paris, http://www.oecd.org/pisa/publications/. [4]

OECD (2019), *TALIS 2018 Results (Volume I): Teachers and School Leaders as Lifelong Learners*, TALIS, OECD Publishing, Paris, https://doi.org/10.1787/1d0bc92a-en. [40]

OECD (2018), *Education at a Glance 2018: OECD Indicators*, OECD Publishing, Paris, https://dx.doi.org/10.1787/eag-2018-en. [3]

OECD (2016), *PISA 2015 Results (Volume I): Excellence and Equity in Education*, PISA, OECD Publishing, Paris, https://dx.doi.org/10.1787/9789264266490-en. [46]

OECD (2016), *PISA 2015 Results (Volume II): Policies and Practices for Successful Schools*, PISA, OECD Publishing, Paris, https://dx.doi.org/10.1787/9789264267510-en. [37]

OECD (2013), *Synergies for Better Learning: An International Perspective on Evaluation and Assessment*, OECD Reviews of Evaluation and Assessment in Education, OECD Publishing, Paris, https://dx.doi.org/10.1787/9789264190658-en. [174]

OECD (2003), *Framework for the Development of OECD Education Indicators*, OECD, Paris, http://dx.doi.org/prova123. [2]

Ollivier, C. (2019), *Enseigner pour Aider à Apprendre les Langues - Approches et Concepts en Didactique de Langue, Théorie et Mise en Oeuvre*, Conseil National d'Évaluation du Système Scolaire (CNESCO), Paris, http://www.cnesco.fr/wp-content/uploads/2019/04/190409_Ollivier-1.pdf. [101]

Pérez Cañado, M. (2016), "From the CLIL Craze to the CLIL Conundrum: Addressing the Current CLIL Controversy", *Bellaterra Journal of Teaching & Learning Language & Literature*, Vol. 9/1, p. 9, http://dx.doi.org/10.5565/rev/jtl3.667. [27]

Peter Dickson and Alister Cumming (eds.) (1996), *Profile of Language Education in 25 Countries. Overview of phase 1 of the IEA Language Education Study*, Nfer, https://www.nfer.ac.uk/media/1452/91128.pdf (accessed on 26 February 2019). [8]

Pfenninger, S. and D. Singleton (2017), *Beyond age effects in instructional L2 learning : revisiting the age factor*, Channel View Publications Ltd, Bristol. [34]

Potowski, K. (2007), *Language and Identity in a Dual Immersion School*, Multilingual Matters, Bristol. [166]

Prapaisit de Segovia, L. and D. Hardison (2008), "Implementing education reform: EFL teachers' perspectives", *ELT Journal*, Vol. 63/2, pp. 154-162. [109]

Pütz, M. (ed.) (1997), *Language Choices: Conditions, Constraints, and Consequences*, John Benjamins Publishing Company, Amsterdam. [90]

Renandya, W. et al. (1999), "A survey of English Language Teaching Trends and Practices in Southeast Asia", *Asian Englishes*, Vol. 2/1, pp. 37-65, http://dx.doi.org/10.1080/13488678.1999.10801018. [122]

Rixon, S. (2017), "The role of early language learning teacher education in turning policy into practice", in Wilden, E. and R. Porsch (eds.), *The professional development of primary EFL teachers. National and international research*, Waxman, Münster, Germany. [110]

Roca Gris, J. (2015), "Categorization of Activities Used in English as a foreign Language: A Corpus Based Study with Pedagogical Materials", *Procedia - Social and Behavioral Sciences*, Vol. 198, pp. 165-173, http://dx.doi.org/10.1016/j.sbspro.2015.07.433. [128]

Rothman, J., J. Cabrelli and K. De Bot (2013), "Third Language Acquisition", in Herschensohn, J. and M. Young-Scholten (eds.), *The Cambridge Handbook of Second Language Acquisition*, Cambridge University Press, Cambridge. [41]

Sánchez Calvo, A. (2007), "A Learner-Centred Approach to the Teaching of English as an L2", *ES: Spanish Journal of English Studies*, Vol. 28, pp. 189-196. [126]

Saville, N. (2012), "Applying a Model for Investigating the Impact of Language Assessment within Educational Contexts: The Cambridge ESOL Approach", *Cambridge ESOL Research Notes*, Vol. 50, pp. 4-8, http://cambridgeenglish.org.ru/images/101052-research-notes-50.pdf#page=36. [156]

Seidlhofer, B. (2011), *Understanding English as a Lingua Franca*, Oxford University Press, Oxford. [91]

Stern, H. (1983), *Fundamental Concepts of Language Teaching: Historical and Interdisciplinary Perspectives on Applied Linguistic Research*, Oxford University Press, Oxford. [114]

Sylvén, L. and P. Sundqvist (2012), "Gaming as extramural English L2 learning and L2 proficiency among young learners", *ReCALL*, Vol. 24/3, pp. 302-321, http://dx.doi.org/10.1017/s095834401200016x. [61]

Taguchi, T., M. Magid and M. Papi (2009), "The L2 Motivational Self System among Japanese, Chinese and Iranian Learners of English: A Comparative Study", in Dörnyei, Z. and E. Ushioda (eds.), *Motivation, Language Identity and the L2 Self*, Multilingual Matters, Bristol, http://dx.doi.org/10.21832/9781847691293-005. [76]

Taylor, J. (2002), "The road to autonomy", *English Teaching professional*, pp. 8-10. [123]

The College Entrance Examination Center (2019), *The College Admission Process in Taiwan*, http://www.ceec.edu.tw/CeecEnglishWeb/E02CEEC.aspx (accessed on 26 July 2019). [154]

Turner, C. and J. Purpura (2015), "Learning-oriented Assessment In the Classroom", in Tsagari, D. and J. Banerjee (eds.), *Handbook of Second Language Assessment*, DeGruyter Mouton, Berlin, Germany/Boston, MA. [119]

Volante, L. and S. Ben Jaafar (2008), "Educational assessment in Canada", *Assessment in Education: Principles, Policy & Practice*, Vol. 15/2, pp. 201-210, http://dx.doi.org/10.1080/09695940802164226. [150]

Wiliam, D. (2011), "What Is Assessment for Learning?", *Studies in Educational Evaluation*, Vol. 37/1, pp. 3-14, http://dx.doi.org/10.1016/j.stueduc.2011.03.001. [158]

Williams, C. (1996), "Secondary education: Teaching in the bilingual situation", in Williams, C., G. Lewis and C. Baker (eds.), *The language policy: Taking stock*, CAI Language Studies Centre, Llangefni, UK. [136]

Yashima, T. (2013), "Imagined L2 Selves and Motivation for Intercultural Communication", in Apple, M., D. Da Silva and T. Fellner (eds.), *Language Learning Motivation in Japan*, Multilingual Matters, Bristol. [96]

Yashima, T. (2002), "Willingness to Communicate in a Second Language: The Japanese EFL Context", *The Modern Language Journal*, Vol. 86/1, pp. 54-66, http://dx.doi.org/10.1111/1540-4781.00136. [95]

Yoshitomi, A. (1990), "Attitudes and Motivation of Japanese High School Students in Learning English as a Foreign Language", *Sophia Linguistica*, Vol. 28, pp. 111-127, http://dept.sophia.ac.jp/is/solific/28_1990/08Yoshitomi.php. [55]

Zheng, Y. and C. Liying (2008), “College English test (CET) in China. Language Testing”, *Language Testing* 25, pp. 408-417, https://www.researchgate.net/publication/234649107_College_English_test_CET_in_China. [148]

Zhou, L. and H. Ito (2011), “English Language Education and Assessment in China”, *Reasearch Bulletin of Naruto University*, Vol. 26. [153]

Annex 5.A. Student learning in programmes in which multiple languages are used for instruction in other subjects

The variety of names to denote education programmes using more than one language for teaching in non-language-related subjects gives an indication of the diverse historical and institutional settings where these programmes have been applied. As far as 5 000 years ago, a foreign language (Sumerian) was used by the Akkadians in the Middle East to learn theology, botany and zoology (Hanesová, 2015[164]; Mehisto, Frigols and Marsh, 2008[140]); in Meiji-era Japan, Kanbun (a form of classical Chinese) was used in private academies to learn history, Confucian ethics and other subjects (Mehl, 2003[165]).

More recently (in the 1960s), so-called "immersion" programmes were created in the French-speaking community of Quebec (Canada) through which, on a voluntarily basis, English-speaking students learnt school subjects in French. The goals were for students who spoke English as their native language to reach a high level of proficiency in French speaking, reading and writing; while reaching normal achievement levels throughout the curriculum (including the English language), and learning to appreciate the traditions and culture of both French- and English-speaking Canadians. This type of programmes became relatively common in the United States as well (Potowski, 2007[166]).

In Europe, Content and Language Integrated Learning (CLIL) started to gain ground after a European Commission's recommendation on learning three languages at school: the native language of the student plus two European languages (European Commission, 1995[167]). In 2006, CLIL type provision was already part of mainstream school education in the great majority of European countries at primary and secondary levels (European Commission, 2006[168]). Following Nikula, Dalton-Puffer and García (2013[169]), the European Commission (2014, p. 3[19]) defines CLIL as "an educational approach in which a foreign language is used as the medium of instruction to teach content subjects for mainstream students".

Even within a country or region and given an overall pedagogical approach, programmes using more than one language for teaching in non-language-related subjects could differ in many ways. For example, they could differ in the languages taught (e.g. French, Chinese, Arabic); in the subjects that are taught through these languages (e.g. history, mathematics); in the overall amount of hours through which the various languages are learnt; and in the characteristics of the students who enrol in them. Given the wide differences across programmes using more than one language for teaching in non-language-related subjects, it does not surprise that the scientific literature does not find an unequivocal answer to the question of how they impact student learning.

Across European countries, whether or not schools offer CLIL programmes was not robustly associated with average school target language proficiency (European Commission, 2012[31]). However, other research suggests that CLIL positively affected spoken production and interaction in Spain (Nieto Moreno de Diezmas, 2016[170]). In addition, some research suggests that the effectiveness of using foreign languages for instruction in other subjects may have a different effect on different foreign language skills. For example, Dallinger et al. (2016[171]) find that German students attending history instruction in English showed greater progress than other students in English listening comprehension, but not in general English skills.

Early research on programmes in which a second language is used for instruction also claimed their potential for equitable education, on the ground that they "open doors on languages for a broader range of learners". However, in a number of educational contexts, researchers observed a strong selection of

students from higher socio-economic background, as well as more motivated and language-proficient students, into these programmes (Bruton, 2013[26]; Pérez Cañado, 2016[27]; Nikula, 2016[23]). In addition, some studies reviewed by Bruton (Bruton, 2013[26]) show that students with learning difficulties are more likely to drop out from this type of education. These phenomena could potentially induce a streaming of motivated and proficient students into bilingual, CLIL or immersion programmes, with an adverse effect on traditional education.

Another concern for policy makers is that these programmes could also potentially harm student learning, as students may fail to understand or make progress with some subject-related content because of the language barriers they face. For example, in 2012, the Malaysian government dropped the ETeMS programme (English for Teaching Mathematics and Science) that it had introduced in 2002 because of concerns that it was negatively affecting mathematics and science learning (whether that was the case is still debated (Nor, Aziz and Jusoff, 2011[172])). Marsh, Hau and Kong (2000[173]) find large negative effects of English-immersion programmes on learning other subjects for upper secondary students in Hong Kong. In contrast, Dallinger et al. (2016[171]) find no relationship between using a foreign language as a medium of instruction and students' learning other subjects, consistently with most literature on this subject (Nikula, 2016[23]).

Annex 5.B. Analyses on the sample of foreign language teachers participating to the OECD TALIS Survey

This chapter presents some results from an analysis on the sample of foreign language lower secondary teachers in the OECD 2018 Teaching and Learning International Survey (TALIS). This sample is composed by teachers who reported to teach modern foreign languages in two different questions of the survey questionnaire.

With reference to the wording of the TALIS questionnaires, the analysis considered foreign language teachers those who:

- Reported that they taught "modern foreign languages (includes languages different from the language of instruction)" to any student in the school and year in which the survey took place
- *And* reported that they taught "modern foreign languages (includes languages different from the language of instruction)" during a particular class chosen from their teaching schedule (the "target class", i.e. the first lower secondary education class they taught in the surveyed school after 11 a.m. on the Tuesday prior to the day they participated in the survey).

The analysis consisted in replicating some indicators published in the *TALIS 2018 Results (Volume I)* (OECD, 2019[40]) and then recalculating these indicators for foreign language teachers. The table/figure number from *TALIS 2018 Results (Volume I)* is always reported in the text. In one case (proportion of teachers in classes smaller than 20 students) the indicator from the *TALIS 2018 Results (Volume I)* has been adapted, so no table or figure number is reported for that specific indicator.

Only some selected results from these analyses are reported in this chapter. The expression "significant differences" refers to differences between foreign language and other teachers that are statistically significant at the 5% confidence level.

Annex 5.C. List of contributing experts for the questionnaires development and their institutional affiliations

Based on the background questionnaires framework, five questionnaires were developed for the PISA 2025 Foreign Language Assessment, which will collect information from students, parents, teachers and at the school and system level. Gabriele Marconi (Analyst, OECD) and Jimena Vargas (Consultant, OECD) developed the questionnaires during 2020 following a process of consultation with a core and an extended group of experts and country delegates. In addition, Etienne Albiser, Francesco Avvisati, Catalina Covacevich, Bruce Golding, Tue Halgreen, Caroline McKeown, Javier Suarez-Alvarez and Choyi Whang (OECD) provided valuable comments.

Core expert group

Jonas Bertling- Educational Testing Service (ETS)

Cédric Brudermann- Assistant Professor of English and Digital Education at the Faculty of Science and Engineering, Sorbonne University

Talia Isaacs- Associate Professor of Applied Linguistics & TESOL, UCL Institute of Education, University College London

Lorena Meckes- Head of International projects at Centro de Medición MIDE-UC, Pontificia Universidad Católica de Chile

Sanneke Schouwstra- Senior research scientist at Cito (National Institute for Educational Measurement)

Extended expert group

Carla Campos- PhD in Didactics of Languages and Cultures, Université Sorbonne Nouvelle Paris 3

Diego Carrasco- Researcher at Centro de Medición MIDE-UC, Pontificia Universidad Católica de Chile

Alister Cumming- Professor Emeritus, Ontario Institute for Studies in Education, University of Toronto

Elsa Fernanda Gonzalez- Professor in the School of Sciences, Education and Humanities, Autonomous University of Tamaulipas

Kathleen Graves- Professor of Education Practice, University of Michigan

Sung Hye Kim- Researcher at the Korea Institute for Curriculum and Evaluation

David Little- Trinity College Dublin, on behalf of the European Centre for Modern Languages

Icy Lee- Professor and Chair of the Department of Curriculum and Instruction, Faculty of Education, the Chinese University of Hong Kong

Lisa Maria Müller- Education Research Manager at Chartered College of Teaching, UK

José Noijons- European Centre for Modern Languages, Council of Europe

Mostafa Papi- Assistant Professor in the School of Teacher Education, Florida State University

6 Reporting proficiency in foreign language skills

This chapter outlines how the results of reading comprehension, listening comprehension and spoken production will be reported in the PISA Foreign Language Assessment. As in other PISA domains, the results will be reported through proficiency scales that can be interpreted in educational policy terms. To facilitate interpretation of the results for educators and policy makers, these scales will be aligned with the Common European Framework of Reference for Languages (CEFR). The CEFR descriptors will be adapted for PISA, and these adapted descriptors are presented.

Reporting proficiency scales

Student performance in the Programme for International Student Assessment (PISA) is reported via proficiency scales that can be interpreted in educational policy terms. For the PISA 2025 Foreign Language Assessment (FLA), three skills[1] will be assessed (reading, listening and speaking) and reported, each on a separate scale. The skills are distinct enough, both logically and empirically, that it is useful to measure them separately (Carroll, 1983[1]).

Information for improving foreign language teaching and learning will result from the analysis of learners' language learning backgrounds and how their backgrounds relate to their proficiency levels, both at the level of individual students (through information collected in background questionnaires) and at the level of entire education systems. Curricular goals indeed are generally articulated in terms of distinct skills, and differences in learners' profiles may be related to differential access to resources, such as native language tutors.

For each skill, PISA will report results on six-point scales that are aligned with the reference levels of the Common European Framework of Reference for Languages (CEFR), originally developed by the Council of Europe (2001[2]; 2020[3]). Reporting results in terms of an existing set of proficiency scales, rather than developing PISA-specific ones, will greatly facilitate the use and correct interpretation of PISA results by educators and policy makers. In addition, by using a reference that is already in use for multiple languages around the world, PISA lays the foundation to expand the range of languages assessed.

Interpreting and using the proficiency scales

The CEFR level descriptors correspond to a progression in language competence within each activity and skill; for each level, a series of "can-do" statements indicate what language learners whose competence falls within that level are typically able to do.

The tasks[2] used in the PISA Foreign Language Assessment to assess reception vary in terms of task requirements and difficulty. Each task provides evidence in support of a particular "can-do" statement in a particular context; together, the tasks are designed to provide reliable evidence on the broad linguistic abilities of students, at the country level.

To ensure adequate scale coverage and targeting of the test, for reading and listening, a preliminary indication of each task's difficulty (based on their conceptual mapping to CEFR can-do statements) is used during the test development phase. After the completion of the test, test responses are analysed using item response theory models. These models allow not only for the combination of students' scores on single items into an overall continuous scale, but also for the comparison of estimates of the overall scores derived from different sets of items on a common scale. These models also allow for the assessment of the reliability of these estimates, as well as the investigation of the presence of bias in particular test tasks against particular groups of test-takers, as defined by country or gender, for example. Item-response-theory models are also used to confirm the relative difficulty of each task: just as students can be ordered from least to most proficient, tasks can be arranged (based on these analyses) from easiest to hardest.[3] Based on this ranking, cut-off scores on the continuous reporting scale can be determined through a standard-setting exercise to construct the six-point scale aligned with the proficiency level descriptors of the CEFR.

Most of the tasks used in the PISA Foreign Language Assessment to assess production do not have an intrinsic degree of difficulty; they are designed to elicit responses that reveal, to an expert rater, the proficiency level of the respondent. Ratings are based on the classification of "speaking" samples into proficiency levels; there is no "correct" or "incorrect" response, but rather, each response may correspond to a particular level of proficiency.

The PISA speaking test is marked by trained assessors using a six-point mark scheme for each part of the test. At each point on the mark scheme, the assessors mark based on a range of linguistic features: discourse management; language resource; pronunciation (the overall focus of pronunciation assessment is intelligibility, regardless of the accent of any particular variety of English). The test results are reported on the CEFR scale that comprises six proficiency levels from Pre-A1 to C1. A candidate's final test score on the speaking test is the average of the scores he or she receives in each part of the test, converted into a CEFR level according to cut-off scores determined by standard-setting exercises.

The term "standard setting" refers to the process used to establish a passing score or grade boundary on an examination, test or practical exam. This score is often referred to as the "cut score". Standard setting requires gathering systematic expert judgement to set pass marks and grade boundaries by looking at test items, tasks and candidate performances.

Common methods used in standard setting are Angoff, modified Angoff, Bookmark and Body of Work. Regardless of the method chosen for standard setting, the series of steps followed are generally similar and contain common elements. Two of the key components of standard-setting activities are representative test content, and a panel of expert judges knowledgeable in the target domain of the content represented on the examination. For the PISA Foreign Language Assessment, the standard-setting exercise will be conducted in accordance with procedures recommended by the Council of Europe's manual on *Relating Language Examinations to the CEFR* (2009[4]).

An overview of English reading, listening and speaking levels for PISA adapted from the CEFR descriptors

In its first foreign language assessment, PISA will assess only English as a foreign language and will not include an assessment of writing. In addition, the assessments of reading, listening and speaking will not cover the full range of situations in which people use these skills. Rather, they will focus on those that are more relevant for 15-year-olds and can be practically assessed within the constraints of PISA.

The preliminary descriptors of proficiency levels were derived by selecting descriptors from the CEFR companion volume (Council of Europe, 2020[3]), which correspond to skills and task situations included in the PISA assessment of English. In particular, the "overall reading comprehension", "overall oral comprehension" and "overall oral production" scales provide relevant descriptors. The original descriptors can be found in Chapter 3, and are only slightly adapted here to fit the particular range of levels and type of tasks included in the PISA 2025 assessment of English as a foreign language (see Table 6.1, Table 6.2 and Table 6.3). In particular, descriptors of levels that are not assessed in PISA have been deleted from the original scales, and other descriptors have been modified to account for the age of candidates, and to either highlight aspects that are explicitly targeted in the assessment or to de-emphasise aspects that cannot be assessed. For more information of the specific components that were deleted or added, see Annex 6.A.

Table 6.1. The PISA scale for reading comprehension in a foreign language

Adapted from the CEFR overall reading comprehension scale

C1 and above	Can understand in detail lengthy, complex texts, whether or not they relate to their own area of speciality, provided he/she can reread difficult sections. Can understand a wide variety of texts including literary writings, newspaper or magazine articles, provided that there are opportunities for re-reading.
B2	Can read with a large degree of independence, adapting style and speed of reading to different texts and purposes, and using appropriate reference sources selectively. Has a broad active reading vocabulary, but may experience some difficulty with low-frequency idioms.
B1	Can read straightforward factual texts on subjects related to likely fields of interest[1] with a satisfactory level of comprehension.
A2	Can understand short, simple texts containing the highest-frequency vocabulary, including a proportion of shared international vocabulary items.
A1	Can understand very short, simple texts a single phrase at a time, picking up familiar names, words and basic phrases and re-reading as required.
Pre-A1	Can recognise familiar words accompanied by pictures, such as a fast-food restaurant menu illustrated with photos or a picture book using familiar vocabulary.

1. This refers to topics empirically established to be typical fields of interest for this cohort age.

Source: Adapted from Council of Europe (2020[3]), *Common European Framework of Reference for Languages: Learning, teaching, assessment - Companion Volume*, © Council of Europe, Strasbourg, p. 54.

Table 6.2. The PISA scale for listening comprehension in a foreign language

Adapted from the CEFR overall reading comprehension scale

C1 and above	Can understand enough to follow extended discourse on abstract and complex topics beyond their own field, though they may need to confirm occasional details, especially if the variety is unfamiliar. Can recognise a wide range of idiomatic expressions and colloquialisms, appreciating register shifts. Can follow extended speech even when it is not clearly structured and when relationships are only implied and not signalled explicitly.
B2	Can understand the main ideas of propositionally and linguistically complex discourse on both concrete and abstract topics delivered in standard language or a familiar variety. Can follow extended discourse and complex lines of argument, provided the topic is reasonably familiar, and the direction of the argument is sign-posted by explicit markers.
B1	Can understand the main points made in clear standard language or a familiar variety on familiar matters regularly encountered at school, leisure etc., including short narratives.
A2	Can understand phrases and expressions related to areas of most immediate priority (e.g. very basic personal and family information, shopping, local geography), provided people articulate clearly and slowly.
A1	Can follow language which is very slow and carefully articulated, with long pauses for them to assimilate meaning. Can recognise concrete information (e.g. places and times) on familiar topics encountered in everyday life, provided it is delivered slowly and clearly.
Pre-A1	Can understand short, very simple questions and statements, provided they are delivered slowly and clearly and accompanied by visuals to support understanding and repeated if necessary. Can recognise everyday, familiar words, provided they are delivered clearly and slowly in a clearly defined, familiar, everyday context. Can recognise numbers, prices, dates and days of the week, provided they are delivered slowly and clearly in a defined, familiar, everyday context.

Note: The 2020 CEFR uses the term “oral comprehension” for this scale, but to simplify this framework and in accordance with the terminology used throughout, the term “listening comprehension” has been used instead.

Source: Adapted from Council of Europe (2020[3]), *Common European Framework of Reference for Languages: Learning, teaching, assessment - Companion Volume*, © Council of Europe, Strasbourg, p. 48.

Table 6.3. The PISA scale for spoken production in a foreign language

Adapted from the CEFR overall spoken production scale

C1 and above	Can give clear, detailed descriptions and presentations on a wide range of subjects, integrating sub-themes, developing particular points and rounding off with an appropriate conclusion.
B2	Can give clear, detailed descriptions and presentations on a wide range of subjects related to a likely field of interest,[1] expanding and supporting ideas with subsidiary points and relevant examples.
B1	Can reasonably fluently sustain a straightforward description of one of a variety of subjects within their likely field of interest, presenting it as a linear sequence of points.
A2	Can give a simple description or presentation of people, daily routines, likes/dislikes etc. as a short series of simple phrases and sentences linked into a list.
A1	Can produce simple, mainly isolated, phrases about people and places.
Pre-A1	Can produce short phrases about themselves, giving basic personal information (e.g. name, address, family, nationality).

1. This refers to topics empirically established to be typical fields of interest for this cohort age.

The 2020 CEFR uses the term "oral production" for this scale, but to simplify this framework and in accordance with the terminology used throughout, the term "spoken production" has been used instead.

Source: Adapted from Council of Europe (2020[3]), *Common European Framework of Reference for Languages: Learning, teaching, assessment - Companion Volume*, © Council of Europe, Strasbourg, p. 62.

Notes

1. Foreign language learners perform communicative language activities (reception, production, interaction and mediation) by using reading, listening, speaking and writing skills (see Chapter 2 for more on these activities and skills). Reading and listening are involved in reception; speaking and writing in production; and all four skills can be involved in interaction and mediation. The PISA 2025 Foreign Language Assessment will assess only reception (reading and listening) and spoken production.

2. In the FLA Framework, "task" is used to refer to any purposeful language activity that language users and learners engage with; this includes, in the context of the foreign language test, any assessment items or exercises (e.g. a set of questions about a text), which are also referred to as (assessment) tasks.

3. For an overview of item-response-theory (IRT) models, see van der Linden and Hambleton (2016[5]). For a review of the use of IRT models in the context of international comparative assessments, see von Davier and Sinharay (2014[6]).

References

Carroll, J. (1983), "Psychometric theory and language testing", in J. W. Oller (ed.), *Issues in language testing research*, Newbury House, Rowley, MA. [1]

Council of Europe (2020), *Common European Framework of Reference for Languages : Learning, teaching, assessment - Companion volume*, Council of Europe, Strasbourg, https://www.coe.int/en/web/education/-/common-european-framework-of-reference-for-languages-learning-teaching-assessment-companion-volume. [3]

Council of Europe (2009), *Relating Language Examinations to the Common European Framework of Reference for Languages: Learning, Teaching, Assessment (CEFR): A Manual*, Council of Europe, Strasbourg. [4]

Council of Europe (2001), *Common European Framework of Reference for Languages. Learning, Teaching, Assessment*, Cambridge University Press, Cambridge, http://www.coe.int/lang-CEFR. [2]

van der Linden, W. and R. Hambleton (2016), *Handbook of Modern Item Response Theory*, Springer, New York. [5]

von Davier, M. and S. Sinharay (2014), "Analytics in international large-scale assessments: Item response theory and population models", in Rutkowski, L., M. von Davier and D. Rutkowski (eds.), *Handbook Of International Large-Scale Assessment: Background, Technical Issues, And Methods Of Data Analysis*, CRC Press, Boca Raton, Florida. [6]

Annex 6.A. Adaptations to CEFR scales

For each table, the information that was deleted for the PISA scale has been crossed out and the information that was added has been highlighted in grey.

Annex Table 6.A.1. Adaptations to the original reading comprehension scale from the CEFR

~~C2~~	~~Can understand virtually all types of texts including abstract, structurally complex, or highly colloquial literary and non-literary writings.~~ ~~Can understand a wide range of long and complex texts, appreciating subtle distinctions of style and implicit as well as explicit meaning.~~
C1 and above	Can understand in detail lengthy, complex texts, whether or not ~~these~~ they relate to their own area of speciality, provided ~~they~~ he/she can reread difficult sections. Can understand a wide variety of texts including literary writings, newspaper or magazine articles, ~~and specialised academic or professional publications,~~ provided that there are opportunities for re-reading ~~and they have access to reference tools.~~
B2	Can read with a large degree of independence, adapting style and speed of reading to different texts and purposes, and using appropriate reference sources selectively. Has a broad active reading vocabulary, but may experience some difficulty with low-frequency idioms.
B1	Can read straightforward factual texts on subjects related to ~~their field of interest~~ likely fields of interest with a satisfactory level of comprehension.
A2	~~Can understand short, simple texts on familiar matters of a concrete type which consist of high frequency everyday or job-related language.~~ Can understand short, simple texts containing the highest-frequency vocabulary, including a proportion of shared international vocabulary items.
A1	Can understand very short, simple texts a single phrase at a time, picking up familiar names, words and basic phrases and re-reading as required.
Pre-A1	Can recognise familiar words accompanied by pictures, such as a fast-food restaurant menu illustrated with photos or a picture book using familiar vocabulary.

Source: Adapted from Council of Europe (2020[3]), *Common European Framework of Reference for Languages: Learning, teaching, assessment - Companion Volume*, © Council of Europe, Strasbourg, p. 54.

Annex Table 6.A.2. Adaptations to the original listening comprehension scale from the CEFR

~~C2~~	~~Can understand with ease virtually any kind of language, whether live or broadcast, delivered at fast natural speed.~~
C1 and above	Can understand enough to follow extended discourse on abstract and complex topics beyond their own field, though they may need to confirm occasional details, especially if the variety is unfamiliar. Can recognise a wide range of idiomatic expressions and colloquialisms, appreciating register shifts. Can follow extended ~~discourse~~ speech even when it is not clearly structured and when relationships are only implied and not signalled explicitly.
B2	~~Can understand standard language or a familiar variety, live or broadcast, on both familiar and unfamiliar topics normally ecountered in personal, social, academic or vocational life. Only extreme [auditory/visual] background noise, inadequate discourse structure and/or idiomatic usage influence the ability to understand.~~ Can understand the main ideas of propositionally and linguistically complex discourse on both concrete and abstract topics delivered in standard language or a familiar variety~~, including technical discussions in their field of specialisation.~~ Can follow extended discourse and complex lines of argument, provided the topic is reasonably familiar, and the direction of the argument is sign-posted by explicit markers.

B1	~~Can understand straightforward factual information about common everyday or job-related topics, identifying both general messages and specific details, provided people articulate clearly in a generally familiar variety.~~ Can understand the main points made in clear standard language or a familiar variety on familiar matters regularly encountered at ~~work~~ school, leisure etc., including short narratives.
A2	~~Can understand enough to be able to meet needs of a concrete type, provided people articulate clearly and slowly.~~ Can understand phrases and expressions related to areas of most immediate priority (e.g. very basic personal and family information, shopping, local geography, ~~employment~~), provided people articulate clearly and slowly.
A1	Can follow language which is very slow and carefully articulated, with long pauses for them to assimilate meaning. Can recognise concrete information (e.g. places and times) on familiar topics encountered in everyday life, provided it is delivered slowly and clearly.
Pre-A1	Can understand short, very simple questions and statements, provided they are delivered slowly and clearly and accompanied by visuals ~~or manual gestures~~ to support understanding and repeated if necessary. Can recognise everyday, familiar words~~/signs~~, provided they are delivered clearly and slowly in a clearly defined, familiar, everyday context. Can recognise numbers, prices, dates and days of the week, provided they are delivered slowly and clearly in a defined, familiar, everyday context.

Note: The 2020 CEFR uses the term "oral comprehension" for this scale, but to simplify this framework and in accordance with the terminology used throughout, the term "listening comprehension" has been used instead.
Source: Adapted from Council of Europe (2020[3]), *Common European Framework of Reference for Languages: Learning, teaching, assessment - Companion Volume*, © Council of Europe, Strasbourg, p. 48.

Annex Table 6.A.3. Adaptations to the original spoken production scale from the CEFR

~~C2~~	~~Can produce clear, smoothly flowing, well-structured discourse with an effective logical structure which helps the recipient to notice and remember significant points.~~
C1 and above	Can give clear, detailed descriptions and presentations on ~~complex~~ a wide range of subjects, integrating sub-themes, developing particular points and rounding off with an appropriate conclusion.
B2	~~Can give clear, systematically developed descriptions and presentations, with appropriate highlighting of significant points, and relevant supporting detail.~~ Can give clear, detailed descriptions and presentations on a wide range of subjects related to ~~their field of interest~~ a likely field of interest, expanding and supporting ideas with subsidiary points and relevant examples.
B1	Can reasonably fluently sustain a straightforward description of one of a variety of subjects within their likely field of interest, presenting it as a linear sequence of points.
A2	Can give a simple description or presentation of people, ~~living or working conditions~~, daily routines, likes/dislikes etc. as a short series of simple phrases and sentences linked into a list.
A1	Can produce simple, mainly isolated, phrases about people and places.
Pre-A1	Can produce short phrases about themselves, giving basic personal information (e.g. name, address, family, nationality).

Note: The 2020 CEFR uses the term "oral production" for this scale, but to simplify this framework and in accordance with the terminology used throughout, the term "spoken production" has been used instead.
Source: Adapted from Council of Europe (2020[3]), *Common European Framework of Reference for Languages: Learning, teaching, assessment-Companion Volume*, © Council of Europe, Strasbourg, p. 62.

Annex A. List of contributing experts and their institutional affiliations

Suzana Bitenc Peharc, Državni Izpitni Center, Slovenia

Letizia Cinganotto, Italian Research Institute for Innovation, Documentation and Educational Research (INDIRE), Italy

Steven Culpepper, University of Illinois at Urbana-Champaign, United States

Alister Cumming, University of Toronto, Canada

Katrijn Denies, KU Leuven (University of Leuven), Belgium

Janet Enever, University of Reading, United Kingdom

Elsa Fernanda González, Autonomous University of Tamaulipas, Mexico

Kathleen Graves, University of Michigan, United States

Talia Isaacs, UCL Institute of Education, University College London, United Kingdom

Haedong Kim, Hankuk University of Foreign Studies, Korea

Kulaporn Hiranbuara, Chulalongkorn University, Thailand

Gisella Langé, Ministry of Education, Italy

Icy Lee, Chinese University of Hong Kong, China

Patsy Lightbown, Concordia University, Canada

Gad Lim, Michigan Language Assessment, United States

David Little, Trinity College Dublin, on behalf of the European Centre for Modern Languages

David Marsh, University of Jyväskylä, Finland

Lorena Meckes, Centro de Medición MIDE-UC, Pontificia Universidad Católica de Chile

Lisa-Maria Müller, The Chartered College of Teaching, United Kingdom

José Noijons, European Centre for Modern Languages, Council of Europe

Gladys Quevedo-Camargo, University of Brasilia, Brazil

Jean Francois Rouet, Centre National de la Recherche Scientifique (CNRS), France

Angeliki Salamoura, University of Cambridge, United Kingdom

Sanneke Schouwstra, Cito (National Institute for Educational Measurement), Netherlands

To find out more about the project, please visit our website at: https://www.oecd.org/pisa/

Write to us

Directorate for Education and Skills-OECD

2 rue André Pascal - 75775 Paris Cedex 16-France

Co-funded by the European Union