

GLOSARIO DE TÉRMINOS RELACIONADOS CON LA EVALUACIÓN (Seleccionados del Glosario ELT de NILE)

LANGUAGE ACQUISITION CONCEPTS

Intake

The language that a learner meets in their environment and that they absorb. A distinction is made between **input** and intake. **Input** is the language available in the environment,

intake is that part of the input that the learner (consciously or unconsciously) chooses to pay attention to and take in. Intake is the first stage in noticing language.

Example "When he hears a foreign language his ears perk up and his eyes brighten-he seems to unconsciously or consciously pay attention to every bit of input that comes his way, busily turning input into intake."

Interlanguage

The version of the target language spoken by a learner at any given time during the period of learning. A learner's interlanguage will change and develop as they become more proficient. Some aspects of it may fossilize as their proficiency develops.

Example "Learners' interlanguage can develop quickly if they get enough exposure – you see the way they use different grammatical structures with more precision, the range and appropriateness of vocabulary use and the clarity of their pronunciation really changing fast."

Receptive / Productive (language skills)

These are terms used in relation to the language skills of reading, listening, writing and speaking. The first two are said to be **receptive** as they involve absorbing language while the latter two are known as **productive** as they involve producing language. Receptive skills are sometimes thought of as being passive while productive skills are thought of as active. In fact, this categorisation is rather misleading, as a reader or listener can be very active in their comprehension and interpretation of language while reading or listening, and of course, much reading and listening takes place interactively with writing and speaking.

Example I think it's rather unhelpful to talk of listening or reading lessons. I prefer to think of integrated skills lessons where a focus on a receptive skill often leads into and supports the learning of productive skill.

ASSESSING CRITERIA

Accuracy

Accuracy describes the ability to write or speak a foreign language without making grammatical, vocabulary, spelling or pronunciation mistakes. It is often contrasted with fluency. Classroom activities are sometimes categorised into those that promote fluency and those that promote accuracy.

Example *"She makes lots of grammar and pronunciation mistakes – her speech isn't very accurate; but she speaks so fluently and expressively that everyone understands her."*

Acquisition

The way in which languages are learnt unconsciously or 'picked up' by exposure to comprehensible input. In this definition, the term *acquisition* is used in contrast to *learning*, which is seen as a deliberate and conscious process of rule learning and self-monitoring of language use. However the terms *acquisition* and *learning* are used interchangeably by some writers.

Example *"She learnt Portuguese simply through acquisition – hearing and reading it all around her and chatting with friends. She never studied it."*

Appropriacy

This refers to the degree of fit or suitability that there is between a piece of language and the social context in which it is used. When the piece of language matches the social context it is said to be appropriate. When it doesn't match it is said to be inappropriate. To match, it needs to be of the equivalent degree of formality. Appropriacy can be seen in pronunciation, vocabulary, grammar or discourse. The terms *appropriacy* and *appropriateness* are often used interchangeably in this meaning.

Example *"I learnt my English by chatting informally with friends. When I started working in an office I had to make a definite effort to get the appropriacy of my language right."*

Coherence

In English language teaching coherence refers to the ways in which a piece of discourse 'makes sense' through links in meaning. It does this by using various internal devices such as

- logical sequencing,
- adherence to a particular genre,
- accepted forms of text structuring,
- accepted external conventions and ways of thinking and experiencing in the outside world, such as
 - adherence to one topic,
 - relevance between topics
 - shared knowledge.

Example -"A: That's the telephone.
- B: I'm in the bath.
- A: O.K." (Widdowson, H. 1978, p. 12)

Cohesion

This is the way in which language is used in written or spoken discourse to make it link together. Cohesion is achieved by using lexical or grammatical devices such as

- lexical fields,
- substitution,
- ellipsis,
- linking words,
- discourse markers,
- back (anaphoric) and forward (cataphoric) reference.

Example "I never understand what he's saying, so I bought a tennis racket." (cohesive, but not coherent sentence)

Communicative competence

Communicative competence refers to an ability to communicate that depends not just on linguistic ability but also sociolinguistic ability, including appropriate use of language, management of discourse and recognising cultural practices in communication e.g. who makes eye contact with who. The growing awareness of communicative as opposed to linguistic competence had a big impact on language teaching and was behind the development of the communicative approach.

Example The use of video in the classroom has made it easier for teachers to focus on communicative competence, as they can show clips in which communication becomes problematic, for example, because the participants don't follow cultural norms for turn taking, or use the wrong register and give offence. Video shows language in context and can lead to awareness and discussion of appropriacy.

Fluency

Fluency is the ability to speak over stretches of language smoothly, naturally and without too much hesitation or pausing. Fluency is sometimes also used to refer to writing. In this case it means writing with ease – coherently and with flow.

Example "He was a native speaker but he spoke so slowly – he was always searching for words, hesitating and pausing. His lack of fluency made him a bit difficult to pay attention to and understand."

Fairness

Children are given enough chances, through being assessed in different ways, to show what they can really do.

Example The teacher uses different ways of assessing children over a period of time, so that she forms a rounded picture of progress and achievement.

Range

Range is a term used in assessment criteria and in syllabus design to refer to the breadth and variety of language (grammar or lexis) that is appropriate for use in a particular genre. For example, the range of language appropriate for use in a text message to a friend about when and where to meet up next is likely to be much narrower than the range needed in a tourist leaflet describing the attractions of an historic town. Teachers are also often encouraged in syllabuses to teach their students the features of an appropriate range of genres.

The semantic range of a word refers to its occurrence across several subsections of a corpus.

Example "I've just marked Pedro's essay – his grammatical range was really quite impressive-he used all the tenses he needed to use, simple and complex sentences, and a variety of conjunctions and discourse markers - just what was needed in that kind of formal essay."

Register

Register has two meanings. It is sometimes used to refer to the type of language (particular vocabulary, grammar or discourse features) that characterises particular fields of language use e.g. nuclear physics, hip hop music, football.

It is also used to refer to the degree of formality of language use, with language generally classified as formal, neutral or informal. The study of register is part of sociolinguistics.

Example "One of the hardest things to learn in a foreign language is using register i.e. what language it is appropriate to use in what context."

TYPES OF TESTS

Achievement test

A type of test which try to measure what students have learnt throughout a course.

It has a direct impact in the classroom context and gives teachers a relevant feedback of the syllabus objectives that have been successfully taught.

Discrete-item tests / Integrative tests

Discrete-item tests focus on eliciting and evaluating parts of language proficiency separately, e.g. grammar, lexis, pronunciation. (Multiple choice grammar ítem tests)

Integrative tests aim to elicit and assess language use as a whole. (Interviews)

Example It is often easier to design and mark discrete-item tests because they focus on just one thing e.g. tenses. Integrative tests, which focus on assessing e.g. learners' ability to speak or write are more complex to mark reliably.

Placement test

A test designed to establish students' previous knowledge according to previously established levels.

It has to be **economic** in the sense of getting the information with the minimum number of items. According to the way it is designed, a Placement test can be **-linear**, where test items follow a pre-established order **or** **-adaptive (CAT: computer adapted test)**, where the items are presented to the test taker according to the answers she gives.

Example OPT (Oxford Placement Test)
OOPT (Oxford Online Placement test) --- an online version of the OPT

Proficiency test

It is a **snapshot measurement of candidates' ability to use a language competence**. It has a predictive validity of the candidates' performance in real life situation

Example A Driving test measures the candidates' ability at a very specific time and predicts how well he/she will do as a future driver.

TYPES OF ASSESSMENT

Analytic / holistic assessment

Two ways of evaluating the performance of learners in order to give grades.

In **analytic** assessment, separate **grades are awarded to different typical features of a performance**.

In **holistic** assessment markers give a grade based on their evaluation of a **learner's overall performance**.

Example When I marked my students' interviews, I did so analytically, giving them a separate mark for fluency, accuracy, discourse management and pronunciation. Later, I discovered that my colleague marked hers holistically, using descriptions of general performance at particular levels. I think I'll try doing that next time, then see which seems better for me and my students.

Assessment for learning

This kind of assessment is often contrasted with *assessment of learning*. It aims at **promoting and encouraging learning rather than just evaluating or assessing it**, seeing assessment as a means of identifying what learning needs to be focussed on

next. It often takes the form of **formative assessment** during lessons and encourages learner autonomy as a way of achieving its purposes.

Example *"Sometimes I video students doing group work, then we evaluate their performance using a checklist. Then together we decide what we need to focus on in the next lessons to help them move forward. This is assessment for learning – they like it and so do I."*

Criterion-referencing assessment

This term refers to a specific way in which **the evidence gathered from an assessment process is used in a way that a candidate's evidence or performance can be referenced or compared with identified criteria**, which define what a learner has to do in order to satisfy the standard.

Unlike a norm referenced test, where your performance is directly related to performances of others, in a criterion referenced test it is your performance alone that matters. If you meet the criteria you pass.

Example The 'Driving Test' is an example of a criterion referenced assessment. The criteria are known before you undertake the test - you must be able to achieve a given percentage of correct answers on a multiple choice 'theory' test; and, then satisfactorily perform a number of practical manoeuvres in the car on the road.

Dynamic assessment

Assessment which **focuses on what the child can do with help rather than what the child can do alone.**

Example A teacher responds to a child's request for help with spellings during a writing task, in the belief that eventually the child will be able to spell the words independently

Formative assessment

Making **judgments about the success of learning while it is taking place rather than once it is over.** The purpose of formative assessment is to help the teacher (or learners) decide what should be taught next, and possibly how, based on analysis of the needs of the learners as revealed by the assessment. Formative assessment is often informal, with the teacher listening to or looking at learners' performance and possibly taking notes. Learners may be unaware that it is taking place.

Example *"Formative assessment really helps me see how well my learners, and individual learners in particular, have learnt something. To help me focus and remember I often use a checklist to monitor them while they are doing groupwork."*

Informal assessment

System for **observation and collection of data about children under normal classroom conditions.**

Example The teacher observes four pairs of children in a problem solving task in class and makes notes on their achievement of the task and the way they interacted

Norm-referencing assessment

This term refers to a specific way in which the evidence gathered from an **assessment process** is used in a way that **a candidate's evidence or performance can be referenced or compared with the performance of others** (other candidates from the same group, for example)

Such an assessment procedure is based on notions that there is a 'normal' curve of distribution for marks and it used to be common in education. However, the assumptions underpinning this approach have been seriously challenged in recent years.

Example The performances of a group of learners are to be grade A to E with an assumption that a pre determined number should get As, Bs, Cs etc.

Objective (assessment)

This term has two main meanings in ELT, one related to assessment and the other to lesson planning.

In relation to **assessment** it refers to types of assessment for which there is only one correct answer and for which the **assessor doesn't therefore need to use their judgment to decide on the value of the answer**. Examples of objective test formats are True/ False, multiple choice, matching, gap-fill.

In relation to **lesson planning**, an objective is **a specification of what a teacher intends the learners to have learnt, or be able to do better, by the end of the lesson**. It is sometimes used interchangeably with *learning outcome* in this meaning.

Example *"The advantage of objective tests is that each item is short and clearly targeted, but their disadvantage is that they don't really test use of the language."*

Portfolio assessment

Portfolio assessment involves the **assessment of a portfolio of work submitted by a learner**. The portfolio may contain compulsory components or be decided on by the learner. The components may include both oral and written work as well as reflections on that work. Assessment criteria are usually used to guide the marking of portfolios so as to stop the marking becoming too subjective.

Example *"For my Spanish course we had to submit a portfolio – I put in it all the reports I'd written as well as corrected versions of them, videos I'd shot as part of my project, and all my project work – questionnaires, tables of findings, photos I'd taken, recordings of interviews. I felt it gave a really rounded view of what my Spanish is like."*

Self-assessment

This is when the learner assesses their own performance, the strategies they have employed to do something or their attitudes. Self-assessment is often a part of formative assessment and is used to enable the learner to become more autonomous in their learning. Self-assessment is often guided by checklists to help learners know what criteria to use for their evaluation.

Example "Students don't always like doing self-assessment at the beginning., but in my experience they get used to it bit by bit and come to see the value of it."

Subjective (assessment)

This term is used in ELT to refer to types of assessment in which the assessor needs to use their judgement as to how correct an answer is, because the answer is open-ended and can be evaluated according to various different criteria. Speaking tests and essays are examples of subjective assessment formats. Two people listening to the same student speaking might grade him/her differently because they are listening for different things or because they give importance to different aspects of speaking.

Example "I was worried about doing an interview as part of my test because I didn't like my grade depending on the examiner's judgment. But then they explained to me that the examiner had to work with specific defined criteria when grading, and that another examiner would be present to grade as well, so I realised my grade wouldn't be subjective."

Summative assessment

The assessment of learning that takes place at the end of a course of learning to see how much of the syllabus covered each learner has learnt.

Example "It's quite difficult to design summative tests – they're meant to reflect what you have taught from the syllabus, but some things are really quite difficult to test, so the test doesn't always reflect the syllabus well enough."

TESTING, ASSESSMENT AND EVALUATION

Assessment and testing

These terms are sometimes used interchangeably to refer to the collection of data about and awarding of marks to learner performance.

-Testing is used just to refer to evaluation involving tests,

-Assessment encompasses not only tests but also other means of assessment such as observation, portfolios, case studies, interviews etc.

Example Some people argue that you get a fairer and more accurate picture of learner performance using the wide range of techniques available through assessment. They think that the results obtained from tests provide a less comprehensive picture of what the learner can do.

Assessment criteria

These are levels or qualities of performance that markers use consciously or unconsciously to grade learners' performance. To prevent assessment criteria being used randomly or unreliably and to guide markers, assessment criteria are very often written out in the form of analytic or holistic (**See analytic/holistic**) band descriptors or checklists.

Example: If you look at http://www.ielts.org/PDF/UOBDS_WritingT2.pdf you will see examples of assessment criteria for writing (Task Response, Coherence and Cohesion, Lexical Resource, Grammatical Range and Accuracy). These have been fleshed out to provide band descriptors for nine levels of language proficiency for IELTS writing.

Evaluation

This is the process of assessing the value of something by collecting data. Evaluation often leads to decision-making. Evaluation can be of teaching, learning, curricula, methods, exam impact, materials or other areas related to teaching and learning.

Example When evaluating materials it is useful to collect not just teachers' opinions but those of learners, too.

ASSESSMENT TECHNIQUES AND CONCEPTS

Benchmarking

It consists of linking assessing materials to predefined standard ones with recognized prestige.

Example Schools have to benchmark their language tests to the CEFR

Descriptive statistics for tests

Whenever we report or analyse the results of a test, there are a number of descriptive statistics which we should provide to give an overview of the distribution of scores on the test.

The **mean** is the average score from all candidates, which is calculated by adding all the total scores on the test and dividing by the number of candidates.

The **mode** is the most frequent score on the test, i.e. the score that most candidates achieved.

The **median** is the score exactly in the middle of the candidates', i.e. 50% scored higher and 50% scored lower.

The **range** is the difference between the lowest score and the highest score + 1 (because in a test with a maximum of 10, for example, you could also score 0).

The **standard deviation** is the average amount that the candidates scores differ from the mean. This is important because it helps describe the shape of the distribution curve.

Item facility refers to how difficult an item is for the sample population. For dichotomous (right / wrong) items, it is the proportion of the group which got it right, e.g. $15 / 20 = 0.75$. Item facility is usually reported as a **p-value** so if 15 out of 20 candidates answered the item correctly, $p = 0.75$

Item discrimination is the second important measure of item analysis. Whereas item facility deals with each item individually, **item discrimination** tells us how the item relates to the test as a whole. Essentially, we want the candidates who get a particular item correct to be the same candidates who score highly on the test as a whole, i.e. stronger candidates should get items right, and weaker candidates should get items wrong.

PARSNIP

This is an acronym for **Politics, Alcohol, Religion, Sex, Narcotics, -isms and Pork**. It refers to the common practice amongst publishers and exam boards of excluding sensitive or **taboo topics** from the content of their products so as not to give offence and to facilitate the sale of these products. Some people believe that this practice is one factor contributing to the lack of real meaning and relevance that is sometimes noted in ELT materials.

Example When you get to know a class, you become aware of their sensitivities and interests. You're then in a good position to judge how much or what parts of PARSNIP to adopt or ignore when choosing materials or topics to use in class.

Portfolio

A portfolio is a collection of a learner's work submitted as a whole and sometimes organised with an index, agreed assignment components and reflection sheets. In ELT, portfolios can contain written work such as essays, emails, reports or video and audio recordings, project work and PowerPoint slides. Portfolios are mainly used for **assessment**. They are also sometimes used in teacher development. A teacher portfolio might contain a CV, some lesson plans, a statement of beliefs about teaching, an action plan, reflections.

Example "An advantage of portfolios is that they allow the learner to express themselves more fully and the teacher to get a fuller idea of a learner's performance than tests can reveal. A disadvantage is that they can take a long time to mark."

Practicality

It refers to the **economy of time, effort and money in testing**. A practical test

- should be easy to design, easy to administer, easy to mark, and easy to interpret the results
- needs to be within the means of financial limitations, appropriate time constraints, easy to administrate, score, and interpret.

Reliability

The consistency of any assessment which means that under the same conditions and with the same student performance the assessment procedure would produce the same results

We talk about **inter-rater reliability** when, using the same assessing criteria, two or more raters can obtain the same results after assessing a student.

We talk about **intra-rater reliability** to check how consistent one rater can be after assessing different students

Example Giving children a multiple choice test in which the two different markers agree on the scores for each child. This may NOT be a valid assessment, of course!

Validity

Where a form of assessment effectiely measures what it intends to measure

Example If children have been reading and writing stories about animals all term, the teacher assesses their work by asking them to write three sentences about their favourite animal

Washback

This is a term (in the US more commonly referred to as **Backwash**) used in testing and assessment to describe the effect on the classroom of tests that the learners will take. Washback may affect e.g. the syllabus, methodology, interaction patterns, attitudes to learning etc., and can be positive or negative.

Example In some countries education authorities deliberately introduce new elements into tests so that they will be used in the classroom. In other words they are relying on the washback effect of a test to bring about change in the classroom. Examples of this might be the introduction of speaking tests or the use of tasks in speaking tests. Washback is sometimes known as 'Backwash' and is contrasted with 'Impact'

TEACHING / LEARNING APPROACHES

EFL / ESL

EFL stands for English as a Foreign Language. Generally speaking, it refers to learners learning English in an environment where English is not used, or to learners

studying English on brief trips to an Anglophone country.

ESL stands for English as a Second Language and has generally been used to refer to learners who have another mother tongue, learning English while living in an English-speaking environment. In the UK nowadays this tends to be called ESOL (English for Speakers of Other Languages). ESOL or ESL classes are likely to include a focus on language and communication, but also on the cultural practices of the Anglophone country the students are living in. With globalisation and the increased movement of people and immigration, the distinction between EFL and ESL is becoming less clear.

Example I teach French in French Guyana where the official language is French. Most of my students speak very little French, though. Their mother tongue might be Portuguese and/or an Indian language. In the street they often hear and speak French Creole. So, am I teaching EFL or ESL?

ELF

This stands for English as a Lingua Franca, and refers to the use of English in international communication. Certain scholars have suggested that as English has become a lingua franca between people from a range of L1s, features of its use such as particular pronunciations and grammatical constructions which would previously have been considered non-standard and 'wrong' should be accepted rather than corrected, providing they do not cause a breakdown in communication, as they are a mark of the L1 learner's identity. There is much debate in ELT about the research base for ELF's findings and their implications for the classroom.

Example If you listened to two non-native speakers of English talking together you might hear them regularly pronouncing the article 'the' as /də/ or /zə/ yet obviously having no problem communicating with one another. ELF proposes that if that's the case there is no need to insist on 'correct' pronunciation with the corresponding loss of learner identity that correction can lead to.

L.O.L.A / L.O.A

LOLA stands for Learning Oriented Language Assessment. It is a systematic approach to language learning that uses formal and informal assessment to:

- help teachers and learners to plan learning more effectively
- measure progress
- identify areas for improvement
- deliver measurable improvements.

Learning Oriented Assessment provides a clear structure for integrating in-course tests, public examinations and less qualitative observations of learners. It helps plan course objectives and to ensure that lessons and study outside the classroom directly contribute to the achievement of each learner's personal objectives.