Listening and reading – towards a delineation of the constructs

Introduction

The purpose of this text is to provide a selective discussion of the constructs of listening and reading, which are labelled as aural and written reception in the CEFR (4.4.2, 65-72). Aspects related to the emergence, functions, forms, activities, processes etc. of listening and reading are addressed. Listening and reading are very complex phenomena and have played and continue to play an important role in individual and social life. In the interest of promoting assessment literacy, this text attempts to provide a selective review of these central language skills. A selective bibliography is also provided to facilitate further exploration. The sources referred to here are also included in this separate bibliography.

The text is organised so that the origins and functions of listening and readings are addressed first. This is followed by a description of the term of construct and its use in language testing and assessment. After that, the listening and reading constructs are dealt with illustrating different needs and different approaches. The CEFR approach to listening and reading is then presented. This is followed by the presentation of the PISA Reading Literacy Framework (originally 2000, and updated several times) as an example of an influential and extensively tested model of assessment of literacy in the mother tongue, which can be consulted for potentially useful information for assessment in foreign languages as well. Some general principles in language testing are then briefly addressed and some concluding remarks are presented.

Listening and reading: origins and functions

One of the obvious questions to ask is: How and when did listening and reading emerge? Another question is: What functions did listening and reading emerge to serve, what functions do they serve at present and what functions might they serve in the future?

The oral modes of language are primary in evolutionary terms – all languages have speech but not all languages even now have writing. The archetype of communication can be taken to be face-to-face interaction. According to Wagner (2014), it is estimated that 50% or more of a person´s time in communicative situations is spent listening.

The invention of writing has indubitably had a profound and lasting influence on individual, social and cultural development. Thus it is not surprising that there was a widespread belief in the mystical powers of writing. For thousands of years, literacy used to be very limited and even today it is not universal. Reading was first reading aloud and reading for a group of listeners. Silent reading (reading without moving lips) emerged much later and first led to great astonishment. Unlike speaking, writing was visually embodied – clay, rock, papyrus, vellum, paper, electronics. As such tools have developed and can be expected to develop, this partly means that written

interaction can be almost as immediate/on line as speaking. Cultural artefacts ranging from the invention of writing a few thousand years ago to modern IT tools have impacted communication modes such that interaction across distance more or less closely can simulate the original pattern of communication.

Writing was originally developed for both mundane needs of increasingly more complex societies (eg. trade across considerable distances) and to conserve holy scripts. Over time, new types of writing emerged to serve new functions (e.g. Britton et.al. 1975; Bühler 1934; Gelb 1951; Jakobson 1960; Kinneavy 1971).

The impact of writing has, thus, been profound and it continues to play the same role as children are helped to gain access to the exciting and rich world of texts, in principle transcending the constraints of place and time. This remarkable achievement is probably not often acknowledged but taken more or less for granted. However, it can be argued that this is an achievement whose fundamental significance is not questioned by any subsequent developments in how people deal with text.

What is a construct?

The term *construct* is addressed briefly next. First, it should be noted that there is no firm agreement on, or definitive definition of, what is meant by the term construct. It is a term which is used in scientific discourse in relation to theories, models, frameworks and similar constructions. A construct appears to refer to a concept/idea developed (inferred, "constructed") to describe and explain an observed phenomenon (eg., language-related behaviour/activity). In testing and assessment, "construct" refers to the concept or the characteristic that a test is designed to measure. The term "construct validity" is used to indicate to what extent an intended characteristic is, indeed, assessed.

In language testing and assessment discourse, constructs became a commonly used term as models of communication and communicative competence emerged. Models of communicative competence have dominated the discussion on constructs since the seminal article by Canale and Swain (1980). Several modifications to their original model have since been proposed (e.g., Bachman, 1990; Bachman & Palmer, 1996; Bacham & Palmer, 2010).

Bachman (2007) describes the different construct approaches in language testing and assessment across almost five decades. He illustrates the long-standing debate on to what extent abilities on the one hand and language use contexts on the other hand are closely related or clearly distinct and affect performance in language assessment tasks. He proposes a distinction between ability-based, task-based and interaction-based approaches to construct definition. Questions about the construct are fundamentally questions about validity.

Six volumes in the Cambridge Language Assessment Series (2000-2006), edited by Alderson and Bachman, represent a skills-based approach to language assessment by surveying the conceptual and theoretical foundations of the specific skills and presenting how they have been and are being interpreted and implemented in assessment

practices.

In language assessment focussed mainly on L1, models with a strong psycholinguistic/cognitive orientation have been proposed drawing on the work on speaking by Levelt. His seminal work on the construct of speaking (1989) was recapitulated in a slightly revised form as a "blueprint of the speaker (1999) and this blueprint was applied to produce a blueprint of the listener (Cutler & Clifton, 1999, see Appendix), of the reader (Perfetti, 1999, see Appendix), and a blueprint of writing (Alamargot and Chanquoy, 2001). They all illustrate the components and processes of the four skills.

One obvious change in the view of constructs in language education and assessment was when the long-lived dichotomy between "active" and "passive" language skills was discarded and replaced by "productive" and "receptive". Developments in cognitive psychology in the 1960s, in particular, had shown that - far from being passive skills - reading and listening are very active processes (being an effort after meaning). The CEFR distinguishes between receptive communicative language activities and strategies, productive communicative language activities and strategies, interactive communicative activities and strategies, and mediating communicative and strategies (57-88).

In a monumental study of the structures of mental constructs, factor analysis has been the standard method (other approaches being multi-trait multi-method and facet theory and mapping sentence advocated by Guttman). Carroll (1993) reanalysed some 1500 published studies of abilities in which correlational and factor analyses had been conducted. He stated that every ability is defined in terms of some kind of performance, or potential for performance at a particular point of time in a class of tasks that vary in difficulty. In Carroll´s view, probably all tasks require more than one ability.

In the 1993 survey he identified a total of 367 abilities in the domain of language behaviour (for instance, 197 factors interpreted as verbal comprehension, 10 as phonetic coding and 15 factors as distinct vocabulary factors (see appendix). Notable more recent contributions are eg., Weir, Vidakovic & Galaczi (2013), Khalifa & Weir (2009; see Appendix), and Field (2013; see Appendix).

Listening and reading, being both internal processes of verbal comprehension, share a number of factors/processes but listening and the testing/assessment of listening is commonly seen to pose specific challenges related to the characteristics of the input (eg., Alderson 2000; Buck 2001; Wagner 2014). As mentioned earlier, the product of listening and reading comprehension (what is comprehended) is gauged by eliciting externalised measures.

Carroll (1976, 37-56) also presented an approach to analyse tests as cognitive tasks. As analytical categories he distinguished the following: stimulus materials (currently usually referred to as input characteristics), overall response to be made at the end of the task, task structure, operations and strategies, temporal aspects of the operation or strategy, and memory store involved. Subsequent models, for instance, Bachman and Palmer´s (1996) framework of task characteristics and target language universe domain (TLU), have applied a similar

approach. Alderson (2001) presents a framework for test design in assessing reading drawing on the TLU approach.

Listening and reading constructs – different needs, different approaches

Reading

It appears that the construct of reading comprehension abilities has received explicit attention only surprisingly recently in testing and assessment - roughly during the past twenty years (Grabe & Jiang 2014; Khalifa & Weir 2009). In his comprehensive volume on assessing reading, Alderson (2001, pp. 116-137) presents a review of recent approaches to defining the construct of reading ability.

The nature of reading comprehension, in particular in L1, has been the subject of continued debates. E.L. Thorndike (1917, 1918) initiated a view of reading as being basically problem solving. This tradition was continued in a revised form displaying a strong trend rooted in early cognitive psychology, which assigned a prominent role to top-down processing, context-based hypothesizing of probable lexical items (utilization of redundancy) and limited eye fixation even of content words (e.g., Smith, 1971). Subsequent developments (e.g., connectionism) challenged the top-down models and pointed out the importance of levels of processing. Thus a bottom-up view gained momentum and efficient word-decoding mechanism, relatively dense eye fixation and rather exhaustive processing and comprehension monitoring were suggested as important processes (e.g., Stanovich, 1992; Kintsch, 1988: construction-integration model, CIA). The debate continues, for instance, on what is the role of the text, of the reader ("the reader variables"; Alderson 2001, 33-60), of the context and of the purpose of reading. As Thorndike proposed, reasoning plays an important role but, according to Stanovich, that role is constrained by a number of other factors.

Synthesising the sources listed above, reading comprehension is commonly considered to require the following *skills and knowledge resources*:

- ability to decode graphic forms for efficient word recognition (phonological, orthographic, morphological and semantic processing)
- ability to access the meaning of a large number of words automatically (vocabulary knowledge)
- ability to draw meaning from the phrase-and clause-level grammatical information (efficient grammatical parsing)
- ability to combine clause-level meanings to build a larger network of meaning relations (discourse-level structuring and genre recognition)
- ability to use a range of reading strategies with more challenging texts (including goal setting, inferencing and monitoring)

- ability to draw on private background knowledge, as appropriate
- abilities to evaluate, integrate, and synthesize information from a text to form a situation model of comprehension (what the reader learns from the text)
- ability to maintain these processes fluently for an extended period of time
- ability to persist in reading and to use the text information appropriately in line with reader goals.

Computers are likely to facilitate some specific types of reading, eg. skimming, reading to search information, and obtaining measures of reading rates and to assess them, for instance for diagnostic purposes.

Listening

As mentioned in the above, it is estimated that 50% or more of a person´s time in communicative situations is spent on listening (Wagner 2014). The important role of listening is also obvious in language acquisition. While listening is a subset of general language ability and it shares many features with reading, there are characteristics that are unique to listening.

Buck (2001) noted that "listening comprehension is a process, a very complex process, and if we want to measure it, we must first understand how that process works". (p. 1) As in the case of reading, listening has been conceptualized both as a mainly bottom-up and a top-down process. Taxonomies of listening distinguish a number of abilities. Valette (1977), based on Bloom´s taxonomy, described a series of increasingly complex cognitive skills that can be used to show increasing facility with listening comprehension: mechanical skills, knowledge of the language, transfer, communication and criticism. Aitken (1978) suggested a communicatively oriented approach with seven component abilities. Richards (1983) lists 33 skills for conversational listening and 18 for academic listening. Weir (1993) distinguished between direct meaning comprehension (4 components), inferred meaning comprehension (4), contributory meaning comprehension (7), and listening and taking notes (2).

Buck discusses the listening construct in detail (2001, 94-115). He identifies the following types of knowledge used in understanding spoken language (the contextualised input to the listener):

- applying knowledge of the language;
- using world knowledge;
- building mental representations of meaning.

His framework for describing listening ability (p. 104) contains four components of language competence (grammatical knowledge, discourse knowledge, pragmatic knowledge, and sociolinguistic knowledge) and strategic competence (cognitive and metacognitive strategies).

According to Buck (2001), it is particularly relevant to the task of assessing listening comprehension to consider

the importance of automatic processing; to take into account the complex interaction of the listener´s knowledge, past experience, current thoughts, feelings, intentions, personality and intelligence; and the rapid processing of language knowledge in real time. This means that listening is a very individual and personal process.
Wagner (2014) claims that assessing a person´s listening skills presents unique challenges. He suggests that presenting spoken texts to listeners raises some questions, eg. should it be presented by a test interlocutor or by a recording? How long should it be, how fast, what kind of language features should it utilize? In short, what characteristics should listening tasks contain.

CEFR approach to listening and reading

The CEFR approach to receptive activities is quoted below in full (2.1.3., p. 14):

- The language learner/user's communicative language competence is activated in the performance of the various ***language activities***, involving **reception**, **production**, **interaction** or **mediation** (in particular interpreting or translating). Each of these types of activity is possible in relation to texts in oral or written form, or both.

- As processes, ***reception*** and ***production*** (oral and/or written) are obviously primary, since both are required for interaction. In this Framework, however, the use of these terms for language activities is confined to the role they play in isolation. Receptive activities include silent reading and following the media. They are also of importance in many forms of learning (understanding course content, consulting textbooks, works of reference and documents). Production activities have an important function in many academic and professional fields (oral presentations, written studies and reports).

- In ***interaction*** at least two individuals participate in an oral and/or written exchange in which production and reception alternate and may in fact overlap in oral communication.

- Not only may two interlocutors be speaking and yet listening to each other simultaneously. Even where turn-taking is strictly respected, the listener is generally already forecasting the remainder of the speaker's message and preparing a response. Learning to interact thus involves more than learning to receive and to produce utterances. High importance is generally attributed to interaction in language use and learning in view of its central role in communication.

- In both the receptive and productive modes, the written and/or oral activities of ***mediation*** make communication possible between persons who are unable, for whatever reason, to communicate with each other directly. Translation or interpretation, a paraphrase, summary or record, provides for a third party a (re)formulation of a source text to which this third party does not have direct access. Mediating language activities – (re)producing an existing text - occupy an important place in the normal linguistic functioning of our societies.

CEFR Illustrative tasks: Reading and Listening

The table below contains a summary of the CEFR text on listening, reading and reception strategies (4.4.2).

Aural reception (listening)	Visual reception (reading)	Reception strategies
The language user as listener receives and processes a spoken input produced by one or more speakers.	The language user as reader receives and processes as input written texts produced by one or more writers.	The language user identifies the context and knowledge of the world relevant to it, activating in the process what are thought to be relevant schemata.
Listening activities include • listening to public announcements (information, instructions, warnings, etc.); • listening to media (radio, TV, recordings, cinema); • listening as a member of a live audience (theatre, public meetings, public lectures, entertainments, etc.); • listening to overheard conversations, etc.	Examples of reading activities include: • reading for general orientation; • reading for information, e.g. using reference works; • reading and following instructions; • reading for pleasure.	Reception strategies include: . Planning: framing (selecting mental set, activating schemata, setting up expectations) . Execution: identifying cues and inferring from them. . Evaluation: hypothesis testing: matching cues to schemata . Repair: revising hypotheses
In each case the user may be listening: • for gist; • for specific information; • for detailed understanding; • for implications, etc. Illustrative scales are provided for: • Overall listening comprehension; • Understanding interaction between native speakers; • Listening as a member of a live audience; • Listening to announcements and instructions; • Listening to audio media and recordings.	The language user may read: • for gist; • for specific information; • for detailed understanding; • for implications, etc. Illustrative scales are provided for: • Overall reading comprehension; • Reading correspondence; • Reading for orientation; • Reading for information and argument; • Reading instructions.	One illustrative scale is provided: Identifying cues and inferring (spoken & written)

In the audio-visual reception, the user simultaneously receives an auditory and a visual input. Such activities include: following a text as it is read aloud; watching TV, video, or a film with subtitles; using new technologies (multi-media, CD ROM, etc.). One illustrative scale is provided for watching TV and film.

PISA Assessment Framework

PISA Reading Framework (2009, 2012) defines reading literacy as follows: "Reading literacy is understanding, using, reflecting on and engaging with written texts, in order to achieve one´s goals, develop one´s knowledge and potential, and participate in society." While PISA (see also PIRLS) deals with the assessment of literacy in L1, its framework provides a detailed and concrete approach which can benefit also in designing assessment in L2 reading.

PISA uses *situation, text and aspect* to organise the domain:

- *Situation* uses the four CEFR domains (personal, public, educational and occupational)
- *Text* uses as text classification the following: medium (print and digital); environment (authored, message-based, mixed); text format (continuous, non-continuous, mixed and multiple); text type (description, narration, exposition, argumentation, instruction – based on Wehrlich 1976 – and transaction)
- *Aspects* are the mental strategies, approaches or purposes that allow teachers to negotiate their way into, around and between texts. Five aspects are distinguished: retrieving information, forming a broad understanding, developing an interpretation, reflecting on and evaluating the content of a text, and reflecting on and evaluating the form of a text. These are also merged into three broader categories: access and retrieve, integrate and interpret, and reflect and evaluate.

PISA also discusses *factors affecting item difficulty*. The following is a citation from the 2012 Framework (p. 69).

The difficulty of any reading literacy task depends on an interaction among several variables. Drawing on Kirsch and Mosenthal´s work (see for example Kirsch, 2001; Kirsch & Mosenthal, 1990), we can manipulate the difficulty of items by applying knowledge of the following aspect and text format variables

- In *access and retrieve* tasks, difficulty is conditioned by the number of pieces of information that the reader needs to locate, by the amount of inference required, by the amount and prominence of competing information, and by the length and complexity of the text.
- In *integrate and interpret* tasks, difficulty is affected by the type of interpretation required (for example, making a comparison is easier than finding a contrast); by the number of pieces of information to be considered; by the degree and prominence of competing information in the text; and by the nature of the text: the less familiar and the more abstract the content and the longer and more complex the text, the more difficult the task is likely to be.
- In *reflect and evaluate* tasks, difficulty is affected by the type of reflection or evaluation required (from least to most difficult, the types of reflection are: connecting; explaining and comparing; hypothesising and evaluating); by the nature of the knowledge that the reader needs to bring to the text (a task is more difficult if the reader needs to draw on narrow, specialised knowledge rather than broad and common knowledge); by the relative abstraction and length of the text; and by the depth of understanding of the text required to complete the task.

In PISA, items that require expert judgment consist of open-constructed and short-constructed responses that requite expert coding. Items that do not require coder judgment consist of multiple choice, complex multiple

choice and closed-constructed response items. The closed-constructed response items are those that require the student to generate a response, but require minimal judgment on the part of the coder.
Short response can be a cross, or circling something. Open-constructed items typically require a whole sentence. Closed-constructed answer; may be writing/copying a name etc from the text

Concluding remarks

Takala, Erickson, Figueras and Gustafsson (2016) note that operationalising any construct definition requires making decisions on formats and methods. Such decisions imply the planning, development and choice of the instruments and the procedures that will be used to collect and analyse evidence. There are several textbooks, manuals and on line materials which address the technical aspects of such implementations and which provide very useful advice on how to approach the complex task.
Takala et al. (2016) point out that there are also Guidelines, Codes of good practice, Codes of ethics and Standards, which address principles such as transparency, accountability, responsibility and fairness and that provide frameworks for the identification and analysis of the implications of the social consequences of test use. Spolsky (2013) presents a good overview on the influence of ethics in language assessment.
Reading and listening are so fundamental components of communication that they are bound to be subject to developments and to remain as topics of continuous exploration and investigation. Some issues for developments in their assessment are:

- How to cover the multicomponent constructs of reading and listening comprehension in a testing/assessment situation?
- To what extent are different texts, tasks, and task types appropriate at different proficiency levels?
- How can/should reading and listening strategies be accommodated in testing/assessment?
- What is the relationship between reading/listening in a testing context vs. reading/listening in non-testing contexts? How can they be brought closer to each other?
- As the reading and listening construct domains are probably underrepresented in all testing/assessment situations, how could construct representativeness be enhanced?
- What are the relative strengths of standardized tests and classroom assessment? How can they be integrated in an optimal way?
- How can integrative tests, in which receptive and productive skills merge, be developed so that their benefits are capitalised on while not risking jeopardising reliability?

If the Council of Europe were to continue developing additional concrete tools for supporting language learning and teaching in Europe, the new tool might focus on integrated testing/assessment and the use of information technology as a medium to facilitate new modes of assessing listening and reading.

Appendix

In this appendix some prominent models of language abilities are presented to allow the users to get an overview of how listening and reading has been modelled.

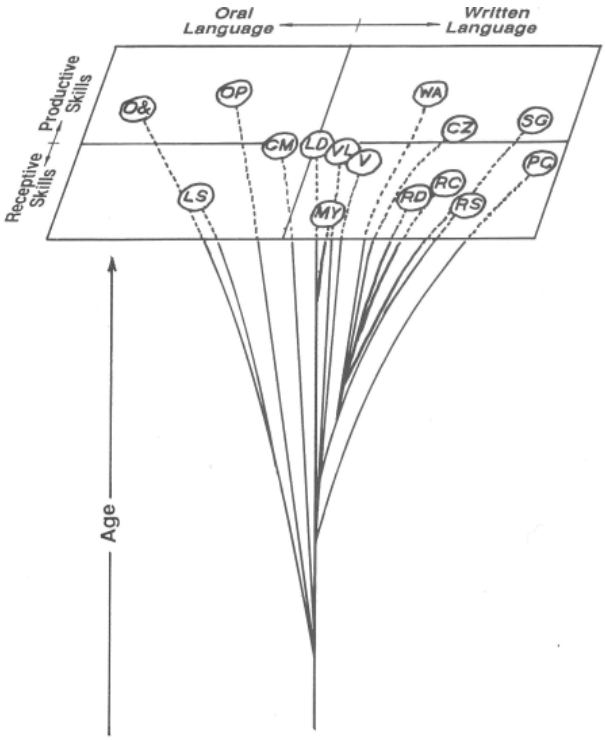
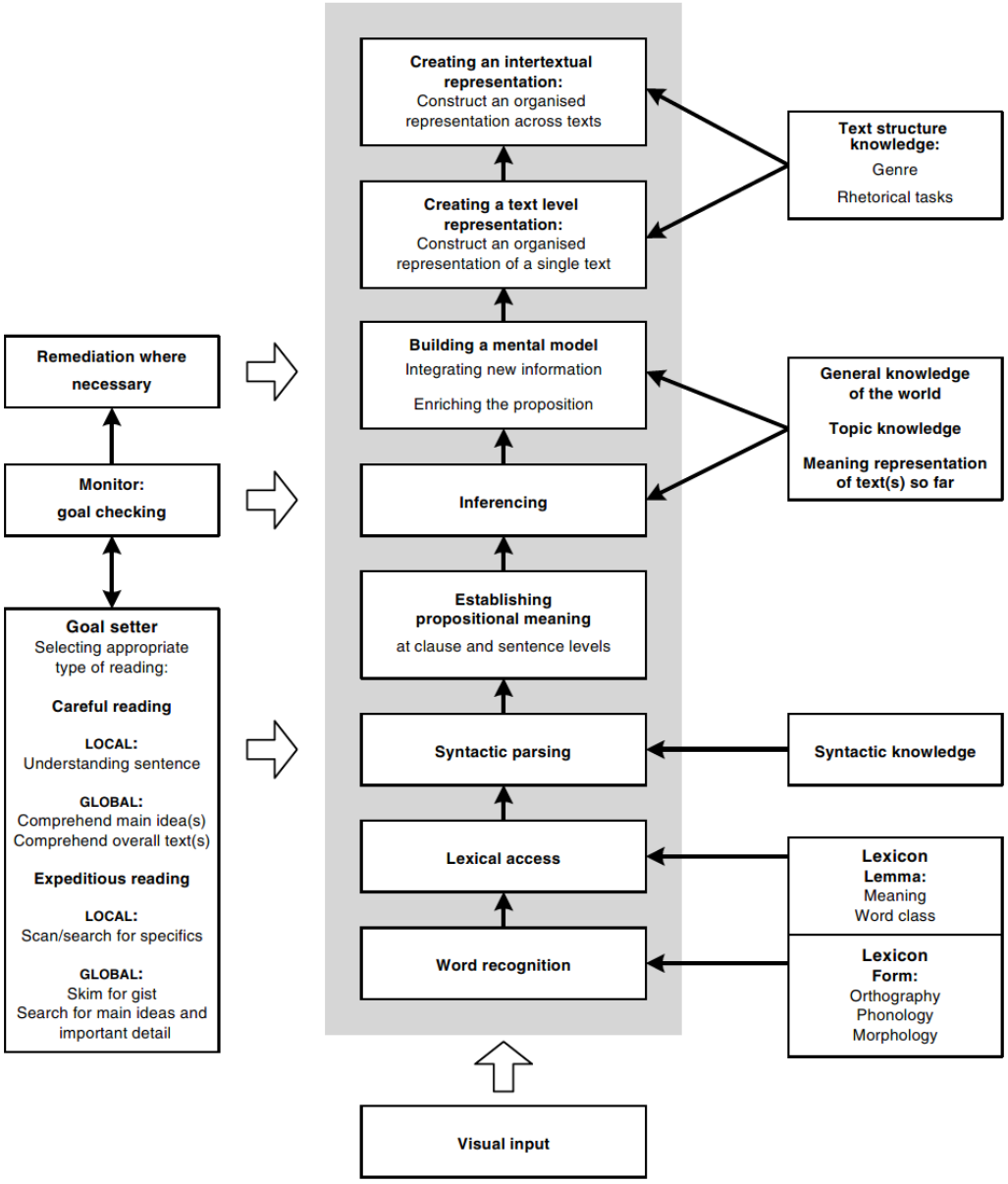


Figure 5.1. Conceptual representation of factors in the language ability domain (see text for explanation).

Conceptual representation of factors in the language ability domain (Carroll, 1990, p. 148)

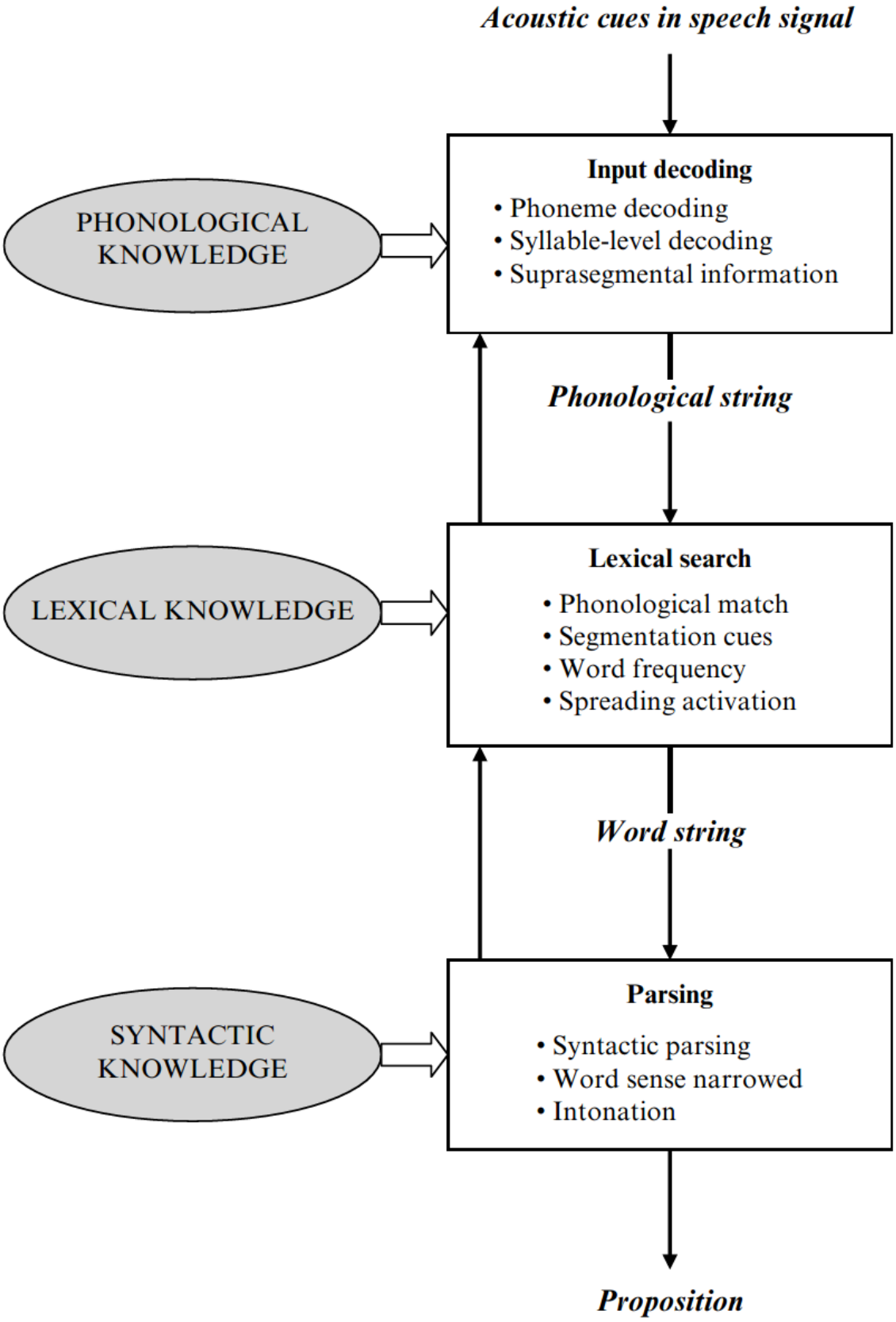
V – verbal or printed language comprehension
LD Language development factors
VL – lexical knowledge factors
RC – reading comprehension factors
RD – reading decoding factors
RS – reading speed factors
CZ – close ability factors
SG – spelling ability factors
PC – phonetic coding factors
MY -grammatical sensitivity factors
LA – foreign language aptitude factors
LS – listening ability factors
CM – communication ability factors
OP – oral production factors
O& - oral style factors
WA - writing ability factors
KL – foreign language proficiency factors

The domain of language behaviour (including 190 factors interpreted as Verbal Comprehension) is composed of 20 factors interpreted as Language Development, 14 interpreted as Spelling Ability, 10 interpreted as Phonetic Coding, 15 factors interpreted as Vocabulary, distinct from Verbal Comprehension (total 367 factors).



Model of reading comprehension by Khalifa & Weir (2009, p. 62)

Model of listening by Field (2013, p. 97)



Blueprint of the reader (Perfetti, 1999)

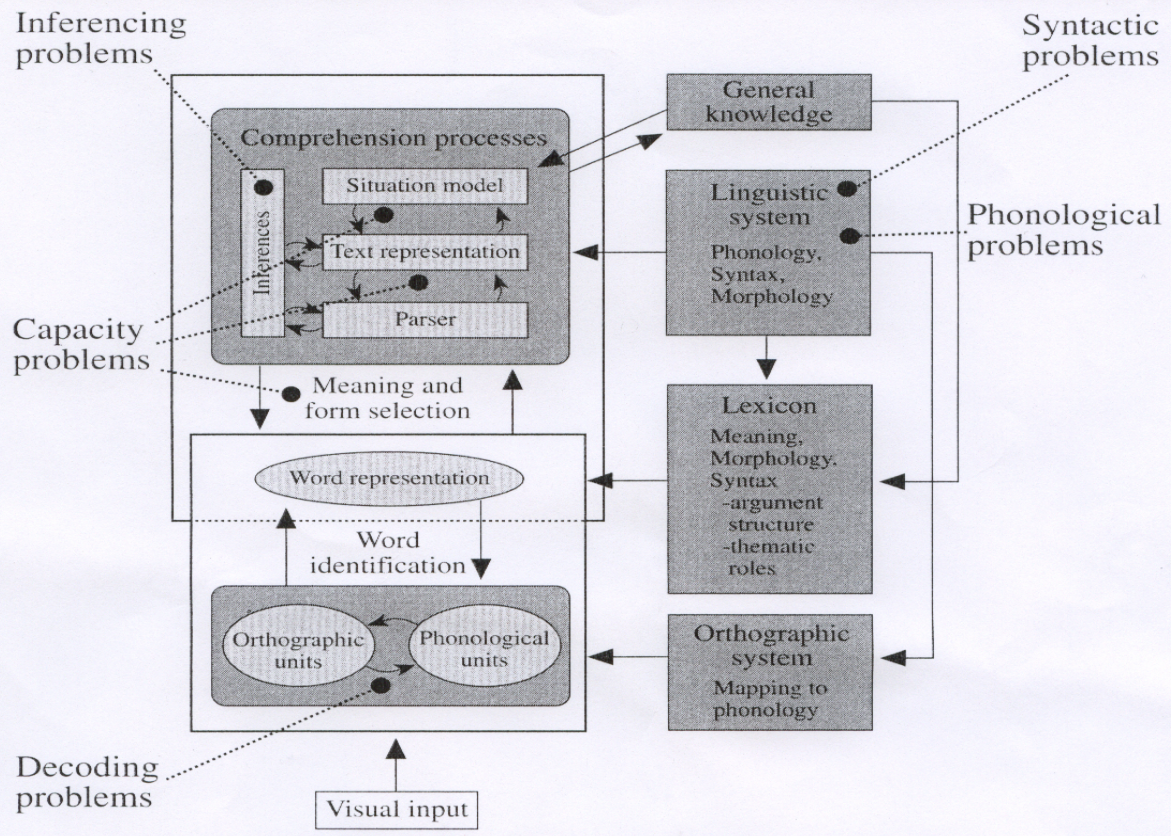
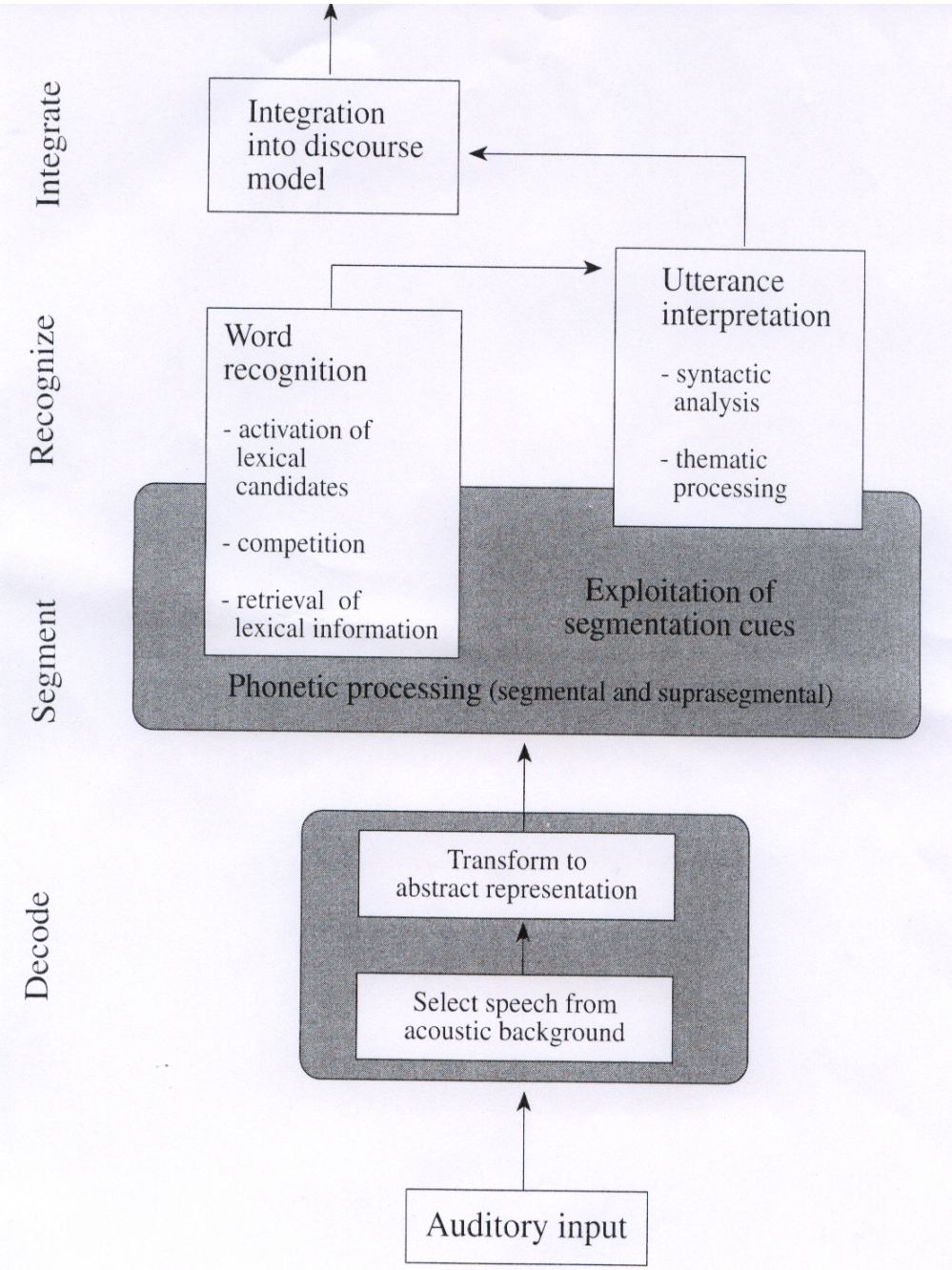


Fig. 6.2 Sources of reading problems. Some components of reading that have been assumed to lead to reading problems. The overall blueprint of the reader from 6.1 is shown in sketch form, with some potential sources of problems indicated. It should not be assumed that the indicated problems are equal in their empirical bases nor that they are mutually independent. Not represented are deficiencies in knowledge, which would have a pervasive negative effect in comprehension.

Blueprint of the listener (Cutler & Clifton, 1999, p. 124)



ig. 5.1 A blueprint of the listener.