doi: http://dx.doi.org/10.5007/1984-8412.2023.e87147

LINGUISTICS MAIN CONTRIBUTIONS TO EARLY LITERACY

PRINCIPAIS CONTRIBUIÇÕES DA LINGUÍSTICA PARA A ALFABETIZAÇÃO

PRINCIPALES CONTRIBUCIONES DE LA LINGÜÍSTICA A LA ALFABETIZACIÓN

Leonor Scliar-Cabral*
Universidade Federal de Santa Catarina

Otília Lizete Heinig**
Universidade Regional de Blumenau

ABSTRACT: Linguistic knowledge is fundamental to the field of early literacy as it underpins methodologies, teacher training, and classroom performance. Therefore, the article's aim is to present the main contributions of linguistics to early literacy. Initially, the linguistic economy principle is examined and how it works, adding to this the contributions of linguistics to the zero or absence concept. Then, the Brazilian Portuguese processing units are defined and presented at the lowest levels, both in the oral and written systems. The difference between phoneme and sound and between letter and grapheme is also detailed to discuss invariance finally. Throughout the text, it is explained how the presented principles can be applied to the early literacy process.
KEYWORDS: Early literacy. Linguistics. Learning. Teaching.

RESUMO: Os conhecimentos linguísticos são fundamentais para o campo da alfabetização, pois fundamentam metodologias, formação de professores e a atuação em sala de aula. Por isso o objetivo do artigo é apresentar as principais contribuições da linguística à alfabetização. Examina-se, inicialmente, o princípio da economia linguística e como ele funciona, somando a isso as contribuições da linguística ao princípio da economia e o conceito de zero ou ausência. Em seguida, são definidas e apresentadas as unidades de processamento do português brasileiro nos níveis mais baixos, tanto no sistema oral, quanto escrito. Detalha-se também a diferença entre fonema e som e letra e grafema para, finalmente discorrer sobre a invariância. Ao longo do texto se explica como os princípios apresentados podem ser aplicados no processo de alfabetização.
PALAVRAS-CHAVE: Alfabetização. Linguística. Aprendizagem. Ensino.

RESUMEN: El conocimiento lingüístico es fundamental en el campo de la lectoescritura,pues sustenta las metodologías, la formación docente y el desempeño de los alumnos. Por lo tanto, el objetivo es presentar las principales contribuciones de la lingüística a la alfabetización. Inicialmente se examina el principio de economía lingüística y su funcionamiento, añadiendo a esto las aportaciones de la lingüística al principio de economía y al concepto de cero o ausencia. Luego, las unidades de procesamiento del portugués brasileño se definen y se presentan los niveles más bajos, tanto en los sistemas orales como escritos. También se detalla la diferencia entre fonema y sonido y entre letra y grafema para finalmente discutir la invariancia. A lo largo del texto se explica cómo se pueden aplicar los principios presentados en el proceso de alfabetización.
PALABRAS CLAVE: Alfabetización. Lingüística. Aprendizaje. Enseñanza.

* Professor Emeritus from the Federal University of Santa Catarina. CNPq. E-mail: leonorsc20@gmail.com.

** PhD in Linguistics from UFSC, retired professor from FURB, teacher trainer and consultant in the area of education and language. E-mail: otilia.heinig@gmail.com

1 INTRODUCTION

The relationship between the alphabetic writing system learning and the theoretical constructs of linguistics is essential to understand that early literacy is a teaching to be developed by those who know the language, its system, and the knowledge object, from the perspective of linguistics, which demands the alphabetic system learning as well as it involves its cognitive and linguistic processes.

Linguistic knowledge for early literacy needs to be understood before any methodology is developed, as it underpins the pedagogical practice and the choices among what to include during each teaching and learning process stage of written language. Due to this edition's main theme, "Early Literacy and Linguistics", we will therefore deal with the contributions of the latter, starting with the linguistic economy concept and how it applies to Saussure's proposal (1972, p. 171) on the structuring of languages, in the paradigmatic and syntagmatic axes, as well as the distributionalism of North American linguistics, pioneered by Bloomfield (1933).

We will continue with the cornerstone in defining the linguistics object (SAUSSURE, 1972), the phoneme definition, whose intuition was the driving force behind the proto-alphabet invention on Mount Sinai. The phoneme definition (JAKOBSON, 1967, JAKOBSON; HALLE, 1971) and its invariant phonetic features (KRUSZEWSKY, 1995; DE COURTENAY, 1972; TRUBETSKOY, 1970) that constitute it, as entities of a psychic nature, that is, representations, differentiate them from the sound that materializes them, of a physical nature and will have profound repercussions on the early literacy methodology and on the respective elaboration of the pedagogical material, since the entity that represents the phoneme, in the alphabetic systems, is the grapheme, also of psychic nature and different from one or two letters, which concretize it, in written BP.

However, a distinction must be made, as invariant phonetic features have the function of distinguishing the meaning between words, while letters' invariant graphic features do not (SCLIAR-CABRAL, 2009)[1]. Another linguistics contributions to early literacy, faithful to the economy principle is that zero or absence have also a value, as long as they are opposed to a sign presence, in the same context (SAUSSURE, 2002, p. 68)[2].

We will also address how languages are structured in our mind at levels, from sensory to cognition and/or, conversely, from cognition to motricity, with the methodological application to early literacy: the processes at the lowest levels have to be automated, such as, in reading, the recognition of invariant features that differentiate the letters from each other, of graphemes and their values and of the phonological word, while, in written production, it is the conversion of phonemes into graphemes, as well as the motor gestures that perform letters, whether cursive or digitized, automations that release creative processes at the highest levels, such as, in reading, the recognition of basic meanings in semantic memory and the attribution of new meanings to words, phrases, clauses, sentences and the macrostructure and, in writing, to planning, which involves the definition of pragmatic intentions, implying what to write and for whom, with the consequent choice of textual types and genre, and how to write (register or style), as well as the chosen words and syntax.

2 THE CONCEPT OF LINGUISTIC ECONOMY AND ITS REPERCUSSIONS

The concept of linguistic economy consists of communicating the maximum, combining the minimum of units in the syntagmatic axis, at a minimum cost in the selection of such units in the paradigmatic axis.

In order to understand this principle, one must take up Saussure's thought (1972, p. 171) and explain how it applies: "The syntagmatic relationship occurs *in praesentia*; it rests on two or more terms equally present in an effective series. On the contrary,

[1] In the case of the Latin alphabet, pioneeringly, Scliar-Cabral inventoried it and discusses it in the article "Recognition of invariances by recycled neurons" (2009).

[2] Le rapport syntagmatique est *in praesentia*; il repose sur deux ou plusieurs termes également présents dans une série effective. Au contraire le rapport associatif unit des termes *in absentia* dans une série mnémonique virtuelle. (SAUSSURE, 1972, p. 171).

the associative relationship unites two terms in absentia in a virtual mnemonic series."[2] That is, syntagmatic relations occur in a chain, combining a unit with the one that comes before and/or after (with the exception of invariant features that do not combine in a chain, but rather simultaneously), while paradigmatic relationships arise from the selection of a unit and only one unit at a time, of the same paradigm in our permanent linguistic memory.

To understand how the principle of economy applies, we must introduce the concept of linguistic levels, which form the oral languages architecture: phonological, morphological, lexical, syntactic, and semantic levels. This architecture exists in all languages. The lowest level is made up of the smallest number of elements, listed in our permanent linguistic memory, as the paradigm of distinctive invariant phonetic features (e.g. [+ or – voice], as in '*casa*' (in Eng. 'house'), / ´kaza/ / '*caça*' (in Eng. 'hunting') / ´kasa/): at the next level, they integrate non-linear, *in praesentia* relationships, to form a very small number of phonemes in each language (eg, the four phonemes /k/, /a/, /z/, /a/). The paradigms of the distinctive invariant phonetic features and the phonemes in each language are in a closed and limited number, resulting in a very high economy and the automation of their processing, both in the reception and in the production of verbal language, as the Phonemic tables 1 and 2 of Portuguese vowels and consonants show.

The methodological consequence, in early literacy, is that the processes at the lowest levels also have to be automated, such as, in reading, the recognition of the traits that differentiate letters from each other, of graphemes and their values, and of the phonological word, while, in written production, is the conversion of phonemes into graphemes, as well as the motor gestures that perform the letters, whether cursive or digitized, automations which release the creative processes at the highest levels, since all the lower levels, with few and closed elements, have to be automated to free the mind for such creative processes as, in reading, the basic meaning recognition in semantic memory and the attribution of new meanings to words, phrases, clauses, periods and the macrostructure. In writing, the creative processes are planning, which involves the definition of pragmatic intentions, implying what to write and for whom, with the consequent choice of textual types and genres, how to write (register or style), as well as the words and syntax.

Phonemes combine to form the 3rd level, which is subdivided (a) into the minimum units endowed with grammatical meaning (also in a closed and limited number, therefore the processes are automatic), e.g. articles, prepositions, suffixes, such as the plural, person and number marks and (b) in the units that refer to meanings external to the grammatical structure, recorded in the phonological mental lexicon of the stems of nouns, verbs, and adjectives.

In Brazilian Portuguese (hereinafter BP), the stems, added by suffixes (including zero) or without them, constitute the syntactic classes. With the syntactic rules (also in a closed and limited number, therefore, the processes are automated), they form the 4th level. It is at this level, in parallel, stress is assigned to the most intense syllable.

Table 1: BP vowel system, according to Quicoli's model (1990), with the addition of nasal vowels

+Orais	**-posterior -arredondado (anteriores)**	**+posterior -arredondado**	**+posterior +arredonda-do**
+alta	i		u
-alta -baixa	e		o
+baixa	ɛ (pé)	a	ɔ(pó)

-Orais (nasais)			
+alta	ĩ		ũ
-alta	ẽ		õ
+baixa		ã	

Source: Scliar-Cabral (2003)

The minimal units, endowed with meaning, are linked to semantic memory (5th level), open to new basic meanings and new semantic fields. This is followed by the increasingly creative and complex levels of meaning construction, resulting from the combinations between the syntactic classes, for the formation of nominal, verbal, and prepositional phrases, clauses, sentences, or periods until reaching the text, from which the macrostructure is extracted.

There is still another parallel level that allows the recognition and production of modalities (declarative, imperative, interrogative, denial, doubt, etc., and their combinations), through morphosyntactic resources, but, above all, through the opposition between intonation patterns. The latter are poorly represented in writing systems by punctuation marks.

Table 2: BP consonants table, according to Lopez (1979), plus glides, examples and comparative terms, following Mattoso Camara Júnior (1953)

		+ant -cor (labial)	+ant +cor (anterior)	-ant +cor (posterior)	-ant -cor -post (posterior)	-ant -cor +post (posterior)
obstruent -cont (stops)	-voiced (voiceless) + voiced	p b	t d			k g (galo)
+cont (fricatives)	-voiced +voiced	f v	s z	ʃ (chá) ʒ (já)		R (rosa)
-obstruinte +nasal		m	n		ɲ (vinho)	
(+vocalic) +lateral -lateral -cons (glides)			 l r (caro)		ʎ (velha) j (pai)	 w (teu)

Source: Scliar-Cabral (2003)

The very architecture of languages and their functioning, which mirrors the structure and functioning of the central nervous system for verbal language, therefore reflects the dichotomy between automatic and creative processes. This principle overturns the arguments of those who condemn phonic methods under the false claim that they force the child to repeat meaningless sounds.

In alphabetic writing systems, there are two more levels. Imagine inviting a friend, with the phrase: "Let's play ball?" One will not think that the air has to resonate entirely through the mouth, in the first 'ball' segment, instead of through the nostrils, so as not to get the word 'mall'! The speaker just quickly and automatically produces the syllable! At this point, he is able to understand that the alphabetic writing systems, as they are secondary to the oral ones, required the introduction of two more levels, precisely the lowest of the alphabetic writing systems, that of invariant graphic features and that of letters.

But how does the economy principle and its repercussions on early literacy apply to all this explanation? It is enough to deduce that, from just eight basic invariant graphic features and a few more for their combinatorics (as we will detail later), it is possible to identify the twenty-six capital letters of the Latin alphabet, of which, in the BP writing system, with one or two of them, all of the less than fifty graphemes that represent the twelve vowels and twenty-one consonants of the BP oral system and that, with these very few elements (as in all written languages that use the alphabet) it is possible to create hundreds of thousands of words: the 2001 Brazilian Houaiss dictionary, for example, registers four hundred thousand words, not forgetting that from the verbs, it only consigns the infinitive and that, continuously, new lexical items are appearing, because the paradigm of ungrammatical words is open to receive new words.

In addition to this potential, the paradigm of basic meanings, in semantic memory, still multiplies it, since the same word can have different meanings in different semantic fields, such as the words 'bridge', 'table' or 'root'. In semantic memory, what we have are the basic meanings, that is, those meanings attributed by the members of the same community to a certain lexical item. In the case of homonyms, in the mental lexicon, there is more than one entry and each one will point to the respective meaning in the semantic memory. On the other hand, when the item is polysemic, the same item will point to the various basic meanings in semantic memory. These meanings, in the enunciation, will be able to refer to present, past or future referents in different contexts.

The repercussions on early literacy are that efficient methodologies must promote creative practices that lead to the recognition automating of which, how many and how the Latin alphabet invariant graphic features are combined (neuronal recycling), as well as of the graphemes and their values (phonemic awareness), for the fast recognition of the written word, a condition for fluent reading and, therefore, for textual comprehension.

This potential multiplies even more when we move to the syntagmatic axis: the combination of words in nominal, verbal or prepositional phrases, these in clauses, clauses in sentences and these in the text lead to the formation of new meanings at each combinatory, a fertile tool for creativity.

A valuable linguistics contribution to the economy principle comes from North American distributionalism (BLOOMFIELD, 1933), since two morphemes whose signifiers are identical (homonyms) can be distinct, by virtue of their distribution, that is, in which grammatical context they appear.

To clarify, some examples from Portuguese are presented, starting with the archiphoneme |S|, whose grapheme is <s>, when it is the suffix or ending of a verb form, as in the examples / ´lavaS/, <lavas> (Eng. (you) wash); /la´vavaS/, <lavavas> (Eng. (you (used) to wash); /lava´raS/, <wash> (Eng. (you) will wash); / ´lavis/, <laves>. (Eng. (He want that you) wash). In all examples, the suffix 's' is marking the second person singular.

Other examples of Portuguese are the unstressed archiphoneme |U(S)| and the also unstressed phoneme /a(S)/, whose graphemes are, respectively, <o(s)>, <a(s)>: they can be an articl (e, as in the examples /aSmI´ninaS/, <as meninas>, (Eng. the girls); a third person personal pronoun, in the oblique, singular or plural form, /kõvi´dejaS/, <convidei-as> Eng. (I) invited them); or a singular or plural demonstrative pronoun /USkIfi´karẽj/, <os que ficarem> (Eng. those who remain).

To conclude the linguistics contributions to the economy principle, we will examine zero or absence concept: they have also a value, as long as they are opposed to the presence of a sign, in the same context (SAUSSURE, 1972, p. 164) and, when in his manuscripts, he states: "nothing has also a value" (SAUSSURE, 2002, p. 68). This principle, when interpreted, underlies the golden rule of

attributing the more salient stress to the written vocabulary pattern of Portuguese: paroxytones words ending with the letters 'a', 'e', 'o', followed or not by 's' (words (clitics), which are very frequent, are excluded, as they do not have a stressed syllable), or ending by 'em', 'ens', 'am', are exempt from the graphic marking stress.

When working with the golden rule of attributing stress to words ending with the letters 'a', 'e', 'o', followed or not by 's', or ending by 'em', 'ens', 'am', we are developing the learner's phonological awareness, as well as when we help him to recognize unstressed monosyllables and disyllables, among those ending in oral vowels, whose graphemes are always written with the letters 'a', 'e', 'o'(s).

Unstressed monosyllables and disyllables, in writing, differ from what happens in speech (glued to the word with a stressed syllable), making its borders opaque and undergoing changes (external closed juncture or sandhi): assimilations, resyllabations, as in the example (adapted, for legibility): *os* + *olhos* = *zoio* (Eng. the + eyes = theys). If, instead of the letters 'e', 'o'(s), 'i', 'u' (s) occur at the end and there is no graphic stress marker, the word is oxytone or stressed monosyllable. Ex.: "João te ama" / "João gosta de ti". (Eng. "John loves me" / "John loves you".

3 DIFFERENCE BETWEEN PHONEME AND SOUND AND BETWEEN GRAPHEME AND LETTER

To assess the linguistics contributions to early literacy, it is necessary to be clear about some essential concepts. Let's start with the phoneme definition: Saussure defines it as an entity of a psychic nature, consistent with the value theory that underlies his proposal, as an oppositional, relative and negative unit (SAUSSURE 1972, p. 164), an object of linguistics study. The phoneme concept underwent developments, as a bundle of distinctive features (JAKOBSON, 1967, JAKOBSON; HALLE, 1971), in which not the phoneme in its entirety, but one or more of its features have the function of distinguishing meanings, and it is to these invariant features, represented in the brain, that the outputs resulting from the various processes through which the acoustic signals captured by the ciliary sensors in the cochlea pass are paired for recognition.

To explain what the archiphoneme is, we have to start from the notion of phoneme: a bundle of distinctive features with the distinguishing meaning function. Thus, in the two words / ˈban/ / / ˈbag/, the [-nasal] feature vs. [+nasal] is what distinguishes the difference in meaning between the two words. It happens that sometimes this function is suspended. This occurs when, for example, due to phonological context conditioning and/or because of sociolinguistic variety, only one of the phonemes is used instead of the other and vice-verse, without changing the meaning.

Let's look at the fricatives /s/ vs. /z/ and /ʃ/ vs. /ʒ/ case of the opposition [-voiced] vs. [+voiced], in the examples in Portuguese: / ˈkasa/, <caça> (Eng. 'hunting') vs. / ˈkaza/, <casa> (Eng. 'house'); / ˈʃa/, <chá> (Eng. 'tea'); / ˈʒa/, <já> (Eng. 'already'). When, however, the position is at the end of a syllable, including a word, the following phenomena will occur: if the following phoneme is [-voiced], that is, voiceless, it is impossible to occur /z/ or /ʒ/ (the same, if followed by silence), as in /mI ˈninUS/, /pURtu ˈgeS/, <meninos>, <português>. But, if the following phoneme is [+voiced], it is impossible to occur /s/ or /ʃ/. We say, then, that the function of distinguishing meaning belonging to the voiced or unvoiced feature has been suspended and we postulate the archiphoneme |S|.

But the suspension may be caused by the articulation point. So, the feature that distinguishes both [+cont] and [cor], that is, /z/ from /ʒ/ or /s/ from /ʃ/, which is [+ or – ant] is suspended at the end of a syllable, including a word, depending on the sociolinguistic variety, since, in such a context, people belonging to to the variety *Carioca*, practiced in Rio de Janeiro will only use the two [-anterior] /ʃ/, /ʒ/ consonants, respectively, in <casta> (Eng, caste) and <mesmo> (Eng, same) and the gaucho will only use the two [+anterior] /s/, /z/ consonants. Note how, in this case, the Portuguese writing system preserved the archiphoneme |S| record, using the same grapheme for all cases: |S| / __# are encoded as <s>, <z>.

With the unstressed vowels /i/ vs. /e/; /u/ vs. /o/ in the final position of a word, a similar phenomenon will occur. According to the sociolinguistic variety, /e/, /o/ will not occur in this phonological context, postulating the archiphoneme |I|, |U|, respectively.

The archiphonemes |U|, |I| are therefore encoded in <o>, <e>. which, in most BP sociolinguistic varieties, result from the neutralization function, respectively, of the phonemes /u/, /o/ / /i/, /e/, in favor of /u/, /i/, when in the final unstressed position of a word, followed or not by the archiphoneme |S|.

Such processes, applied to early literacy, lead to the inference that we are facing a competitive context, as the respective graphemes can be either <u> (marked form), <o> (*defaut* form)/ <i> (marked form), <e> (*defaut* form), but the marked form is very rare and is usually followed by the grapheme <s>, in addition to the fact that they are paroxytone words, whose most intense vowel is coded, obligatorily with a graphic stressed mark, as in <ônus>, <lápis>. What constitutes one of the biggest problems for the learner is that, in most sociolinguistic varieties, the realization of the archiphonemes |U|, |I| is more frequent, respectively, in unstressed [u], [i] sounds, causing a conflict between them and the coding of archiphonemes in <o>, <e>.

Letter is not synonymous with grapheme: the recognition of their invariant features does not depend on the language that adopts a system, e.g. Latin, so the processes to recognize them and the methods for the early literacy student to automate it are the same for any learner whose cultures adopt the same system.

For example, the word "passa" (Eng. "pass") has five letters in written Portuguese and four in written Engliss "pass", but only four graphemes in Portuguese, <p>, <a>, <ss>, <a>, each representing, respectively, the phonemes /p/, /a/, /s/, /a/ and only three graphemes in English, <p>, <a>, <ss>, each representing, respectively, the phonemes /p/, /ɛ/, /s/. Note that if both the grapheme and the phoneme have the function of distinguishing meanings, they can only be taught when extracted from the context of words: for example, if I replace <p> with <m>, or /p/ with / m/, I will get the words <massa> (Eng. <pasta>), in the written system and / ´masa/, in the oral system. If I replace the third sound of the word "carta", spoken by a caipira (it is a retroflex [ɽ]), by the third sound of the same word "carta" spoken by a gaucho from the border (it is a multiple vibrant alveolar [r]), the sounds are completely different, yet it is not the realization of distinct phonemes, but the same archiphoneme |R|, as there was no change in meaning. These are phonetic variants (sounds), or realizations of the same phoneme or archiphonemes, the latter being psychic entities of an abstract nature.

It is therefore imperative to be clear about the difference between letter and grapheme: the last is a unit that, in a given written language, has the function of distinguishing meanings and representing phonemes. In <date> the letters are the same for Portuguese and English, but not graphemes. In BP, <d> represents / ´d/; <a>, /a/; <t>, /t/; <e>, |I|; in Eng., <d> represents / ´d/; <a>, /ei/; <t>, /t/; <e>, zero. The letters are the same, but the graphemes are not, because the phonological system is unique to each oral language.

Developing phonemic awareness does not consist in asking the child to produce isolated sounds (sometimes impossible to pronounce without vowel support such as stops [p], [b], [t]. [d], [k], [g]) and, yes, those who perform the phonemes of a given language, since the aim is to make the child manage to concretize the phonemes, producing the sound, while tracing with the finger one or two letters that concretize the graphemes that represent phonemes.

It is also important to distinguish between knowledge of phonemes for use, spontaneously acquired during language acquisition, which every speaker has of his language and conscious knowledge, which is learned during correct early literacy.

It is clear that the goal of automating letters tracing recognition and the relationship between graphemes and phonemes is to obtain a very fast written word recognition, so that the individual can read fluently and thus understand what he reads. Next, we will explain another concept from linguistics, that of invariance.

4 INVARIANCE

Although the term invariance has been used more recently by cognitive linguistics in the semantic sense attributed to it by Lakoff (1990) that the topological properties of entities in the source domain of the metaphor are mapped as properties of the corresponding entities in the target domain (BRUGMAN, 1990, p. 257), or by Chomsky (1968, p. 68-69), with regard to universal

syntactic complexity, in this chapter, the focus is on the concept arising from linguistics, applied to the invariant features that distinguish capital letters of the Latin alphabet.

The concept of invariance espoused here goes back to that of phoneme, already mentioned. The substance/form dichotomy, developed by glossematics (HJELMSLEV, 1953), was applied to all linguistic levels: phonological, morphological, lexical, syntactic and semantic. This dichotomy already appeared in Saussure's (1972, p. 27) speech circuit theory: acoustic waves (substance of a physical nature) are the physics or acoustic phonetics study object and articulated gestures, the object of physiology, or articulatory phonetics, while the linguistics object, such as, for example, the phoneme are, for Saussure (1972, p. 164), entities of a psychic nature.

Neuroscience, in particular, using magnetoencephalography (MEG), associated with behavioral experiments, empirically proves that recognition takes place by pairing a few pertinent traits with those stored in permanent memory, as demonstrated by Vogels and Biederman (2002).), including in monkeys, as mentioned by Dehaene (2012, p.158): "In collaboration with the neurophysiologist Rufi Vogels, Irving Biederman also showed that numerous neurons in the inferior temporal cortex of the monkey resisted metric deformations, preserving their non-accidental properties" and, further on (DEHAENE, 2012, p.159): "Complex objects are recognized mainly from the configurations that form their contours. The junction points between these contours, which form the T, L, and E visual configurations." The evidence brought by neuroscience confirms, then, what was theoretically already postulated by linguistics: it is evident that we will not be able to have registered with all the details, in our mind, the phenomena that we will face, in order to recognize them. Let us, therefore, examine the invariances involved in the recognition of the written word.

The first of the invariances in written word recognition is the perceptual one: changes in the size of the letters do not interfere with its identification. The same bottle size letter in a book cover title is recognized as the same that appears in size 12 on the catalog card the title. What is important is to recognize the invariant features, sometimes very tenuous, that differentiate the letters from each other, although, in each letter, in the same stroke, the differences in size are relevant. Thus, in the capital letter E, the vertical line is always a little longer than the three horizontal lines.

The second invariance is that of letter position. In this case, it is necessary to be very careful and, only by understanding the difference between letter and grapheme, will we understand what letter position invariance is. For example, to recognize that the letter **e** is the same in the first and second syllables of the word **pele**, it is independent of the position in which it is in the word, as it is recognized by the two invariant features that compose it **c**, -, not by the position it occupies in the word. But as graphemes that have the function of distinguishing the meaning of written words, they represent distinct phonemes: the first <e> represents the lowest anterior oral vowel and the second, the archiphoneme |I|, precisely because it is in the final unstressed position of the word, therefore, their values depend on their position in the word. As can be seen, despite all stylistic variation, the invariant features that distinguish the letters from each other were preserved, sometimes only one, as in the *uelo* example in which the letter *u* is only different from *o* because the circle is not complete.

The recognition of the invariant features, of the letters does not depend on the infinite variants in their execution, such as size, case (UPPERCASE or lowercase), font and style (printed, handwriting, *italicized*, **bold** or <u>underlined</u>…), or the position they occupy in the word (DEHAENE, 2012, p. 33-34), details discarded for recognizing which letter it is.

Some experiments that demonstrate what is really relevant for letter recognition are those by Dehaene *et al.* (2002), when presenting the first word subliminally (29 milliseconds): the effect on the target word was the same, whether both were written in the same font or not, with a reduction in activity in the left ventral occipital-temporal region. However, it is observed that this effect does not occur in the primary visual region, since it is sensitive to changing sources. Only the left ventral occipital-temporal region operates with font invariance, that is, with more abstract constructs, crucial for attributing the same values to graphemes, even when in different fonts the same letters do not share any trait, as for example, A/a, G/g, M/m, which is why the right hemisphere recognizes them as different from FAIR/fair (DEHAENE *et al*., 2004).

The same occurs with children who have not become literate and who only recognize logos such as Coca-Cola: only the region in the right hemisphere lights up during the experiments and not the left ventral occipital-temporal region.

We will present, in the sequence, the invariant features that differentiate the capital letters.

5 LATIN ALPHABET PRINT CAPITAL LETTERS INVARIANT FEATURES

The lowest level of reading processing, properly speaking, is the letters invariant features recognition (in this case, the Latin print alphabet). The primary invariant features of such letters are eight: **| O ɿ c U ⊃ ~ .**

More traits are added:

- The straight line position: vertical, horizontal or inclined (on the stick, only on the letter **y**): | \ - (Examples, **I V A Á À**); **ɿ** (**n y**).
Quantity of each stroke: 1, 2, 3, 4, 5 or 6 (Examples, **I L Z F E É Ê**).
- Different size in the same font: the straight one (Examples, **F**).
- Exceeding the imaginary baseline (lowercase only): **g j p q y.**
- Direction and how they are combined: to the right of the axis (**b**); to the left of the axis (**d**); vertex down (**V**); vertex up: (**A**); stick with opening at the top, to the left: (**a**); stick with opening at the top, to the right: (**f**); stick with opening at the base, to the right: (**t**) ; stick with opening at the base, to the left: (**J**); semicircle opening to the right: (**q**) ; semicircle or half ellipse opening to the left: (**b P**);
- Topological combinations: small line at the top, either cutting a third of the cane, or the base of the circle, or making an angle with the semicircle: (**r t Q G**).

None of the graphic features has the function of distinguishing meanings (unlike the invariant acoustic features of phonemes).

We reiterate that letter is not synonymous with grapheme: the recognition of their invariant features does not depend on the language that adopts a system, for example, Latin, so the processes to recognize them and the methods for the early literacy student to automate it are the same for any learner whose cultures adopt the same system.

The greatest difficulty in learning to recognize the letters invariant features resides in the fact that vision neurons have not been programmed to recognize the differences in the features direction, which is essential to distinguish several letters from each other, such as b / d. Therefore, an epigenetic change is required, with the systematic teaching of neurons, called by Dehaene, 'neuronal recycling': (DEHAENE, 2012, p. 166).

6 FINAL CONSIDERATIONS

In this article, we presented the main contributions of linguistics to early literacy. Initially, we examined the linguistic economy principle and how it works, thanks to the way in which languages are structured along the syntagmatic and paradigmatic axes, in an architecture made up of lower levels, starting with invariant traits, whose operation needs to be automated and at different levels. increasingly creative and complex.

In addition, we defined and presented the BP processing units at the lowest levels, both in the oral and written system, namely, the invariant phonetic features, phonemes and archiphonemes, as well as the invariant graphic features and the capital letters of the Latin alphabet and written Portuguese graphemes. We also detailed the concepts of linguistic zero or absence and invariance and we demonstrated how these principles apply to quality early literacy.

We approached all these aspects, because we consider that, for a quality early literacy, which allows the learning of the alphabetic writing system in less time, it is necessary the teacher being endowed with solid linguistic knowledge. These deal with verbal language, its acquisition, learning and processing, that involve sciences such as linguistics, psycholinguistics, neuroscience and neuropsychology.

With this knowledge, the early literacy teacher will have the necessary theoretical bases to choose the best methodology and which pedagogical materials are suitable for teaching work with their students. Consequently, it will help students to make them able to critically understand texts that circulate socially, as well as to produce texts that meet their objectives, in school, professional or personal environments.

REFERENCES

BLOOMFIELD, L. *Language*. New York: Henry Holt, 1933.

BRUGMAN, C. What is the Invariance Hypothesis? *Cognitive Linguistics*, v.1, n.1, p. 257–266, 1990.

CHOMSKY, N. *Language and Mind*. San Diego: Harcourt Brace Jovanovich, 1968.

DE COURTENAY, J. N. B. *A Baudouin de Courtenay anthology*: The beginnings of structural linguistics. Bloomington, In: Indiana University Press, 1972.

DEHAENE, S. *Os neurônios da leitura*. Porto Alegre: Penso, 2012.

DEHAENE, S. *et al*. Cerebral mechanisms of word masking and unconscious repetition priming. *Nat. Neurosci*, v. 4, n. 7, p. 752-758, 2002.

DEHAENE, S.; JOBERT, A.; NACCACHE, L.; CIUCIU, Ph.; POLINE, J.-B.; LE BIHAN, D. e COHEN, L. Letter binding and invariant recognition of masked words: behavioral and neuroimaging evidence. *PsycholSci*, v. 15, n. 5, p. 307-313, 2004.

DEHAENE, S. *et al*. Cerebral mechanisms of word masking and unconscious repetition priming. *Nat. Neurosci*, v. 4, n. 7, p. 752-758, 2002.

HJELMSLEV, L. *Prolegomena to a theory of language*. Trad. Francis J. Whitfield. Baltimore: Waverly Press, 1953.

JAKOBSON, R. *Fonema e fonologia*. Tradução: J. Mattoso Camara Jr. Rio de Janeiro: Acadêmica, 1967.

JAKOBSON, R.; HALLE, M. *Fundamentals of language*. The Hague: Mouton, 1971.

KRUSZEWSKY, M. *Writings in general linguistics*. Amsterdam: John Benjamins Publishing, 1995.

LAKOFF, G. The Invariance Hypothesis: is abstract reason based on image-schemas? *Cognitive Linguistics*, v.1, n.1, p. 39–74, 1990.

LOPEZ, B. A. *The sound pattern of Brazilian Portuguese* (Cariocan dialect). Tese de doutorado. Los Angeles: Univ. of California, 1979.

MATTOSO CAMARA Jr., J. Para o estudo da fonêmica portuguêsa. Rio de Janeiro: Simões, 1953.

QUICOLI, A. C. Harmony, loqweringandnasalization in Brazilian Portuguese. *Lingua*. Aalborg, Denmark: Elsevere, n. 80, p. 290-1990.

SAUSSURE, F. de. *Écrits de linguistique générale par Ferdinand de Saussure*. Texto estabelecido e editado por S. Bouquet & R. Engler. Paris: Gallimard, 2002.

SAUSSURE, F. de. *Cours de linguistique générale*. Edição crítica preparada por T. de Mauro. Paris: Payot, 1972.

SCLIAR-CABRAL, L. Reconhecimento das invariâncias por neurônios reciclados. *Signo*, v. 34, n. 57, p. 2-14, 2009.

SCLIAR-CABRAL, L. *Princípios do sistema alfabético do português do Brasil.* São Paulo: Contexto, 2003.

TRUBETSKOY, N. S. *Principes de phonologie*. Trad. de J. Cantineau. Paris: C. Klincksiech, 1970.

VOGELS, R.; BIEDERMANN, I. Effects of illumination intensity and direction on object coding in macocaque inferior temporal cortex. *Cerebral Cortex*, v.12, p. 756-766, 2002.

Received on April 25, 2022. Approved on May 21, 2022.