

Speech processing, human mind and literacy: how social diversity affects the engineering of educative games

Luiz Antonio G. Senna
Graduate Program in Education
University of Rio de Janeiro State
Brazil
senna@senna.pro.br

Abstract: This paper is concerned to some specific theoretic issues subjacent to the usage of games into literacy practices, more specifically those not assisted in locus by any literacy agent, in which speech processing figures as a necessary user/game interface tool for pedagogical purposes. It describes the motivations of a particular approach of *phonemic representation*, based on a theoretic conception of mental unity called *macrophoneme* which suffers straight influence from social experience. The text is segmented into three parts, respectively: main contextual motivations of the usage of certain games with speech processing tools for literacy purposes; a brief description of the theoretic mental model used to explain what macrophonemes are and how they are represented; some recommendations about the usage of this kind of entity as conceptual unities in games engineering process.

This paper is concerned to some specific theoretic issues subjacent to the usage of games into literacy practices, more specifically those not assisted in locus by any literacy agent, in which speech processing figures as a necessary user/game interface tool for pedagogical purposes. It describes the motivations of a particular approach of *phonemic representation*, based on a theoretic conception of mental unity called *macrophoneme* which suffers straight influence from social experience. The text is segmented into three parts, respectively: main contextual motivations of the usage of certain games with speech processing tools for literacy purposes; a brief description of the theoretic mental model used to explain what macrophonemes are and how they are represented; some recommendations about the usage of macrophonemes as conceptual unities in games engineering process.

Main context & object definitions

Global demands of access to literacy: despite of all official efforts word wide, writing still remains as one of the most determinant indicators of social exclusion; absent of writing skills may be more intense among underdeveloped countries, but also figures inside peripheral communities in the developed world (UN 2010); digital technology tools can contribute to minimize, both, economic and didactic costs within the development of public policies of mass education concerned to literacy, mostly where traditional schools infrastructure is not available or not sufficient; the usage of digital technology into literacy practices involves, although, some constraints concerning the type of instruments applicable and the concepts of grammar unities under which written language is understood.

Games as strategy for the contextualization of writing skills: the usage of written languages, as it's mental representation, does not attend the same structural basis applicable to oral languages (SENNA 2010a); oral languages, so called *natural languages*, are product of a phylogenetic predisposition, a human universal property; written languages are cultural instruments, private to those who belong to one same cultural community and, at least, structurally defined upon the kind of comprehension each culture has about oral language structure; that's why there are so many distinct types of writing systems, as f.eg., alphabetic, left-to-right or right-to-left oriented, ideographic or iconic, or so many types of alphabets.

Most cases of analphabetism or severe costs in literacy process are identified among individuals from the periphery of the hegemonic culture that determines the official writing system to be learned and reproduced in civil life; for those individuals, literacy begins when written language is presented as a material possibility into their lives, something that can belong to them as an instrument of daily life (Boto 2004), not as an instrument of domination; the usage of games oriented to literacy purposes can reduce the cost of self identity of peripheral communities in face of written language and its particular skills; games are organized under very well defined rules whose motivations are locally recognized as part of a playful effort of interaction (Grigorowitschs 2010); games contribute to the development of social-affective concepts subjacent to written language mental representation; digital games may also contribute into mass education systems, through distance education facilities or supplementary formal education resources.

Speech processing and games for literacy: once applied to literacy, and aiming peripheral communities, games must attend three major principles, that may be taken as *parameters of adequacy in terms of social interests*: (i) keep users attention towards the usage of written language to represent oral language, (ii) sustain a playful effort of interaction in face of any oral input offered by user and (iii) present an orthographic form for any lexical demand set by user during the game experiencing; as a condition for social adequacy, there might be also considered two technical properties, so considered parameters of adequacy in terms of potential for dealing with social diversity in the interface game-users: (i) game must be open to any speech input, what assures its capacity of interacting with any kind of cultural community, even those from minorities, and; (ii) must be able to recognize distinct sets of phonemic patterns inside the range of sets of one same language system, what assure its capacity of processing any dialects, linguistic registers and ordinary fluctuations proper of language natural derive (Lucchesi 2004).

These two technical parameters – absolute necessary for the adequacy of games oriented to literacy, in the contexts of social diversity – are tied up, although, to constraints on the primary definition of *phoneme*, the unity subjacent to any conceptual effort associated to speech.

Phonemics and social diversity in the mind

Phonemic representation and the foundations of Science: Phonemics (also known as Phonology) is the field of Linguistics dedicated to the study of the phonic structure of natural languages, by means of their physical vocal production, or their mental representation into a symbolic concept known as *phoneme*; there are several hypothesis and theories about the mental interface that links vocalizations to mental representations and vice-versa; however, the more phonemic theories vary from each other, there's one aspect that remains intact and common to all of them: the classic, Cartesian, single and distinctive conception of a *phonemic unity* in the basis of mental representation of language sound pattern; like almost everything in the so called sciences of language, the basic patterns of all theories come from the assumption of a fix human representation, considered universal and whose theoretic structures delineated what is still considered as “*speaker*”, that abstract, universally conceived entity (Carston 1988).

The conception of this entity, for its turn, derives from a prior assumption about a fix, also universal conceived, mental model, which came to us since the Antique Greece, passed through Descartes and Kant, and reached us coming direct from Piaget and Noam Chomsky: *the occidental model of mind*; this theoretic model is, at the same time, product and inspiration of the singular mathematic way of understanding the world, through which – along History – we all learned to distinguish what's right and wrong, and also to comprehend what does it mean thinking (Piaget 1983).

Based on that occidental model of mind, phonemes used to be defined as a set of distinctive unities, equally represented in the minds of all speakers under the domain of one same natural language, what – according to this hypothesis – grants all those speakers can understand each other while users of a same grammar system; phonological variations are, then, understood as part of the dynamic structure which systematically interferes on each phonemic entity on a regular basis, resulting the same variations for any speakers of the same language or dialect; this conception of variation is based on the principle that any phoneme may vary on a predictable range of phonic acts, but still remains the same mental entity in any case; that's why phonemic description enrolls the

identification of the whole set of phonemes of a natural language and, therefore, the description of the allowed variations each phoneme may suffer in the usage of that particular language (Bybee 2001).

Language mental representation may, however, be analyzed from the point of view of another type of mental model, not ruled by the occidental patterns of knowledge construction, what would bring strong consequences to the ordinary concept of phoneme, and, for extension, to the engineering of educative games oriented to literacy practices.

Mind, cultural diversity and subjectivity of human's representations: L. Vygotsky's *Social Formation of Mind*, first published 1934, is the first organized theory concerned to a non-Cartesian model of mind; not exactly oriented to the structural dimension of cognition, Vygotsky's social mind proposed a new interpretation of humans capacity of representation, based on the absence of any previous, innate knowledge that could provide meaning for the things we live by, into a cultural, local and historic set of semantic features; basis of the very early conception of a pragmatic approach to a mental model, Vygotsky open new frontiers to the understanding of such a mind that could sustain so much diversity among humans, spite of one universal cognitive structure; this approach resulted revolutionary – still nowadays – once brought new lights on the controversial relation between structural backgrounds of mind, so called innate basis of cognition, and semantics of ordinary representations, usually told to belong to humans subjectivity, far from any classic reasoning; Vygotsky remarked an ontological indissociation between structural and subjective representations inside human's mind.

Compared to this social model of mind, Bergson's theory about humans representation process could be finally analyzed from another point of view: in human's representation, time gaps from one conceptual state to another, as from one arrange of impulses to another, may vary, both, from culture to culture, and/or from subjectivity to subjectivity; so, the limits and characters of any mental concept may be restrict to the single mind by which they have been processed; denying the comprehension of perceptual activity as neutral in terms of subjectivity, Vygotsky admits that perceptual activity doesn't result universal representations, but non-neutral, subjective, meanings and concepts; truth or non-truth reasoning come into the sphere of *adequate* or *non-adequate* judgments, derived from the confrontation of one's self meanings and/or concepts to another one's meanings and/or concepts, what is concerned to the usage of the one device of knowledge construction in Vygotsky's mind, called: *the proximal development zone*.

Phonemes and mental diversity: spite of a good lack of controversy among linguists, almost all of them agree with phonemes being associate to some perceptual activity and further mental representation; figure 1 present a chart of all possible vowels one can identify in any natural language, according to the International Phonetics Association; the area demarked inside red square corresponds to the sequence of vocalizations from the tonal pattern phoneme /a/ to the topical anterior phoneme /i/, passing, respectively, through /e/ and /e/ positions (as in “*b-ed*”, open, and “*be-long*”, close); each one of the dots (“•” in figure 1) represents the very typical form of each phoneme, which is obtained when all oral cavity variables (buccal walls and mouth) reach certain typical positioning; like a violin's neck, mouth's positioning for each phoneme in a sequence from /a/ to /i/ is not oriented by any physical mark, so that the dots distributed in figure 1 are absolutely abstract representations of ideal placements, located somewhere variable from mouth to mouth; individual singularities concerned to the mobility of mouth parts, such as more/less tongue distension or tonicity, interferes in the amplitude range of that line from /a/ to /i/, or may produce distinct dots positioning; in consequence, final vowel representations inside a user's mind will inexorably result his own vowel representations, substantively distinct from another ones; the shorter the line from /a/ to /i/, the less distinctive will be those four phonemes from each other inside the users mind, and vice-versa; once each one's experience of language acquisition depends on the sonorous material heard from adults linguistics utterances and also from the vocalizations produced by himself, it is absolutely unreasonable to propose a theory to defend that the phonemes of a language may be considered one single thing in the mind of all those who speak that language; this diversity of phonemic representations, from mind to mind, should never be absorbed into any theory based on the classic conception of phoneme, associated to a neutral conception of human's perception and representation.

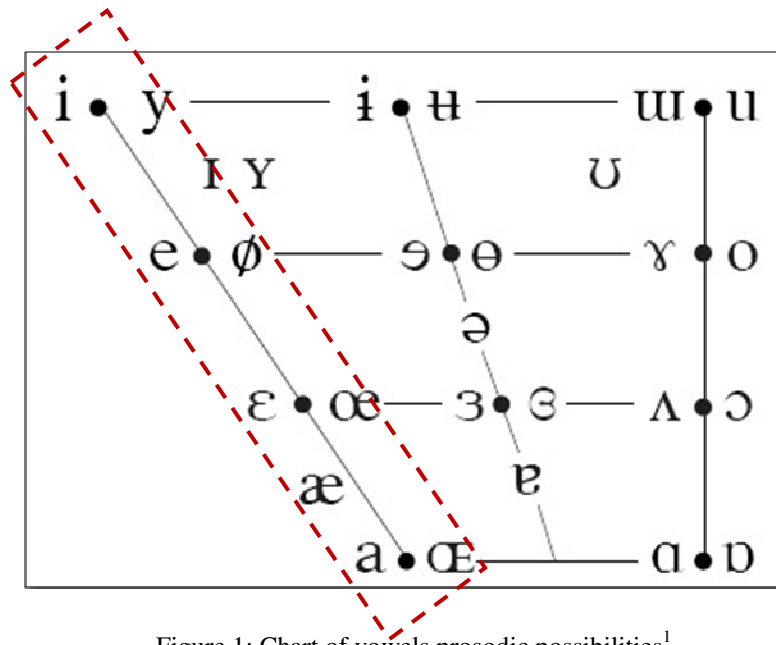


Figure 1: Chart of vowels prosodic possibilities¹

The diversity of phonemic mental entities may be theoretically solved by the assumption that humans *proximal development zone* derives cognitive conditions to deal with phoneme as a complex entity, open to a range of actualizations, from a typical point to unlimited others, located on the frontiers where the distinctivity face to another phonemes will be not clear out of a pragmatic context of language usage; this complex range of possibilities all speakers of a same language identify as one single phonemic entity is called *macrophoneme* and may be considered the basis of a whole theory of grammar whose operational unity is a metaphoric category (Senna 2010a).

Phonemic representation as support for games modeling

Automatic speech processing: speech processing belongs to Artificial Intelligence and is defined as a robotic function, consisting of three major activities: detection of vocal input, analysis of vocal input properties and lexical recognition; early speech processing theories were based on classic phonemic parameters (Chomsky 1957), taken from an essentialist approach of phoneme physical properties, so called *phonological features*; those first attempts already reported the plasticity of phonemes in current oral speech, so that they were defined as a set of phonological features which suffer systematic transformations depending of each phonetic context of an oral utterance form; this plasticity should be explained and described by a set of rules called *phonological transformations*; by the end of last century, artificial intelligence derived a new approach of vocal representation, no more based on classic phoneme, but on a new parameter of mathematic unity, absolutely plastic thou still conceived as a set of physical features; this unity derived from *Parametric Theory*, that brought the perceptual notion of *typicality* to define unity as a branch of infinite actualizations around a conceptual *type*, varying from 100% typical to gradually less typical; Parametrics demanded the development of a whole theory of pattern classification (Duda et al. 2001), much more inspired on Physics than on classic Phonology; since then, speech processing has been developed as a robotic approach to the acoustics of oral speech’s features and their behavior face to typicality parameters (Huang et al. 2001; Alencar & Alcaim 2008); robotics, however, imposes constraints to the usage of speech processing routines, mostly concerned to the amplitude of semantics interface between recognized material and meaning.

Meaning is a huge problem when the task is human communication: most of meaningful unities into a communicative act derives from non logical operations, such as inferences, metaphors or ironies; once almost all

¹ From **The International Phonetic Alphabet** (ver. 2005) app. “The International Phonetics Association” at: <http://www.langsci.ucl.ac.uk/ipa/vowels.html>. Visited Oct. 10 2010.

speech processing applications are nowadays concerned to cell phones users interface or commercial strict purposes automation on costumer services, restrictions on the semantics of speech processing are solved by the usage of close phrase patterns to which recognized vocal inputs are compared (Quateri 2001).

Speech processing and games oriented to literacy: according to the adequacy parameter previously defined here, game applications oriented to literacy must be open to distinct cultural systems and to distinct vocal inputs of a same language; once, as argued before, macrophonemes define sets of phonological patterns that are true just inside one's mind, no previous set of parameters should allow robots to identify words from the recognized vocal inputs, without an implicit knowledge about general laws that (a) restricts the amplitude and (b) rules the parameter settings concerned to phonemic mental representation; as a matter of fact, literacy games propose another further theoretical challenge, namely, (c) the adequate set of parameters that explain what goes inside one's mind when select a written alphabetical form to represent an oral utterance; (c) is a determinant knowledge for the usage of a literacy game in distance education or non-assisted activity, because it allows the robot to understand what kind of comprehension about alphabetic system does the user present and what kind of parameters user's mind apply for defining his own macrophonemes (Senna 2010b).

Conclusions: The usage of games in literacy practices may contribute to minimize costs of access to written language among individuals coming from the periphery of modern culture; however, due to the nature of written systems, games must be sufficient open to cultural diversity and able to simulate real communication acts, so to contribute to the development of a sentiment of belonging to written practices; speech processing assumes a very central role in the engineering of such games, but it is necessary to enlarge actual robotic capacities of associating recognized vocal inputs to semantics; macrophonemes may contribute to the development of new robotic routines of speech processing, by presenting a new reasoning about mental parametrization process.

References:

- ALENCAR, V; ALCAIM, A (2008). *Atributos eficientes em reconhecimento automático de voz distribuído*. In: **SBA: Controle & automação** Sociedade Brasileira de Automática. Vol. 19(2). ISSN 0103-1759. Natal/RN. PP: 147-154.
- BERGSON, H (1896) **Matter and memory**. Available at www.forgottenbooks.org (2010).
- BOTO, C (2004) *Aprender a ler entre cartilhas: civilidade, civilização e civismo pelas lentes do livro didático*. In: **Educação e Pesquisa**, São Paulo, v. 30, n. 3, Dec. 2004 . Disponível em <http://www.scielo.br/scielo.php?script=sci_arttext&pid=S1517-97022004000300009&lng=en&nrm=iso>. Acessado em 08 Jan. 2010. doi: 10.1590/S1517-97022004000300009.
- BYBEE, J (2001). **Phonology and language use**. Cambridge/UK: CUP.
- CARSTON, R (1988) *Language and cognition*. In: NEWMeyer, F (Ed.) **Linguistics – the Cambridge Survey**. Vol. 1. Cambridge/UK: Cambridge University Press. Pp: 38-68.
- CHOMSKY, N; HALLE, M (1957) **The sound pattern of English**. Cambridge/MS: MIT Press.
- DUDA, R O; HART, P E; STORK, D G (2001) **Pattern classification**. New York: John Wiley & Sons.
- GRIGOROWITSCHS, T (2010). *Jogo, mimese e infância: o papel do jogar infantil nos processos de construção do self*. In: **Revista Brasileira de Educação**. v. 15(44). PP: 230-246.

- HUANG, X; ACERD, A; HON, H (2001). **Spoken language processing** – a guide to theory, algorithm and system development. Upper Saddle River/NJ: Prentice Hall.
- INTERNATIONAL PHONETICS ASSOCIATION (1999) **Handbook of the International Phonetic Association: A Guide to the Use of the International Phonetic Alphabet**. ISBN: 978-0521637510. Cambridge/UK: Cambridge University Press.
- LACEY, H (2008) *Aspectos cognitivos e sociais das práticas científicas*. In: **Sci. stud.**, São Paulo, v. 6, n. 1, Mar. 2008 . Disponível em <http://www.scielo.br/scielo.php?script=sci_arttext&pid=S1678-31662008000100004&lng=en&nrm=iso>. Acessado em 08 Jan. 2010. doi: 10.1590/S1678-31662008000100004.
- LUCCHESI, Dante (2004) **Sistema, mudança e linguagem** – um percurso na história da linguística moderna. São Paulo: Parábola.
- MORENO, Montserrat ET AL. (1998) **Conhecimento e mudança** – os modelos organizadores na construção do conhecimento. Versão Portuguesa: São Paulo: Moderna: 2000.
- PIAGET, J (1970) **A Construção do Real na Criança**. Rio de Janeiro: Zahar.
- PIAGET, J (1971) **A Epistemologia Genética**. Petrópolis: Vozes.
- PIAGET, J (1983) “Psicogênese dos conhecimentos e seu significado epistemológico”. In: PIATELI-PALMARINI, M (1983) **Teorias da linguagem e teorias da aprendizagem** - o debate entre Jean Piaget e Noam Chomsky. São Paulo: Cultrix. PP: 39-76.
- QUATERI, T (2001). **Speech signal processing** – principles and practices. Upper Saddle River/NJ: Prentice Hall.
- RABINER, L R; JUANG, B H (1993). **Fundamentals of speech recognition**. Upper Saddle River/NJ: Prentice Hall.
- RABINER, L R; SCHAFER, R W (1978). **Digital processing of speech signals**. Upper Saddle River/NJ: Prentice Hall.
- SANTANA, S M; ROAZZI, A; DIAS, M G (2006) *Paradigmas do desenvolvimento cognitivo: uma breve retrospectiva*. In: **Estud. psicol. (Natal)**, Natal, v. 11, n. 1, Apr. 2006 . Disponível em <http://www.scielo.br/scielo.php?script=sci_arttext&pid=S1413-294X2006000100009&lng=en&nrm=iso>. Acessado em 08 Jan. 2010. doi: 10.1590/S1413-294X2006000100009.
- SENNA, L A G (2010a). *Psicolinguística e sistemas gramaticais no campo da Psicopedagogia*”. In: GODOY, H & SENNA, L A G. **Psicolinguística**. Curitiba/PR: IBPEX.
- SENNA, L. A. G. (2006) *Categorias e sistemas metafóricos - um estudo sobre a pesquisa etnográfica*. In: **Educação em Foco** (Juiz de Fora), v. 11(1), p. 169-188.
- SENNA, L. A. G. (2007) *O conceito de letramento e a teoria da gramática - uma vinculação necessária para o diálogo entre as ciências da linguagem e a educação*. In: **DELTA** - Documentação de Estudos em Linguística Teórica e Aplicada, v. 23, p. 45-70.
- SENNA, L. A. G. (2007, Org.) **Letramento – princípios e processos**. Curitiba: IBPEX.

SENN, L. A. G. (2010b) *Onde está o meu aluno nas teorias sobre alfabetização? Reflexões sobre as ausências no caminho entre a teoria e a prática de letramento em EJA*. In: COSTA, R. e CALHÁU, M.S. **E uma educação pro povo, tem?**, Rio de Janeiro: CAETES. p. 47-64.

UNITED NATIONS (2010) **United Nation literacy decade: education for all**. Imprint document: 66th General Assembly. A/C.3/65/L.9/Rev.1. New York: UN.

VYGOTSKY, L (1987) **Pensamento e linguagem**. SP: Martins Fontes.

VYGOTSKY, L S (s/d) **A formação social da mente**. São Paulo: Nacional. 2007.