Repercussions of the Inability of Organic Scripts on Second-Language Acquisition

Jayme Sanches Filho[1]

April 18th, 2024

There are no conflicts of interest to disclose.

Abstract

This work postulates the thesis of the Inability of Organic Scripts (IOS), which argues that the writing systems that evolved naturally, shaped exclusively by sociocultural factors, are inherently unable to adequately serve second-language acquisition (SLA) because their limitations interfere deeply and negatively with the process, pushing the foundations of the target L2 (second language) outside of the learner's zone of proximal development during a crucial stage of SLA. This work ultimately advocates for the elaboration and usage of Linguistically Expanded Scripts to help overcome the obstacles caused by IOS during SLA.

Keywords

Inability of organic scripts — Linguistically expanded script — Second-language acquisition — Zone of proximal development

Introduction

SLA (second-language acquisition) is the process through which a second language (L2) is acquired. The term also refers to the studies of how a second language is learned; that is, to the research of how learners come to internalize the linguistic system of a second language and how they make use of that system during speech production and comprehension (VanPatten and Benati 2010, 2). As an intrinsically interdisciplinary field, SLA impacts and draws from many different areas (e.g., linguistics, psychology, psycholinguistics, sociology, discourse analysis, conversational analysis) and employs varied methodologies to understand the process of language acquisition (Gass and Selinker 2001, 1; Robinson, Sawyer, and Ross 2001, 12; Juffs 2011, 277).

The present work argues that the naturally-evolving writing systems used

1 Independent researcher. ORCID: 0009-0002-8250-5880. Correspondence: *jssfo@protonmail.com*.

around the world—here referred to as *"organic scripts"*—carry important and varied limitations that render them unable to serve as adequate media for SLA. The root of these limitations lies in the lack of scientific insight that permeated said writing systems' origins as well as development, which were guided almost exclusively by sociocultural and political factors.

The assortment of said limitations is here referred to as the Inability of Organic Scripts (IOS). Of great importance to this work is the hypothesis that the IOS leads to either the emergence or aggravation of a wide range of obstacles during SLA.

Also important to this work is Vygotsky's concept of *zone of proximal development* (ZPD), which describes the range of tasks learners can perform with adequate mediation of a more knowledgeable other (e.g., a skilled instructor, a more knowledgeable peer), as opposed to the range of tasks learners can already perform on their own and the range of tasks learners cannot yet perform even with mediation (Kurt 2020), respectively referred to as *zone of actual development* (ZAD) and *zone of distal development* (ZDD).

This work argues that the main way in which IOS hinders SLA is by pushing foundations of the target L2 out of learners' ZPD during a pivotal phase of the acquisition process. This is because IOS creates unnecessary obstacles during the initial stages of SLA that cannot be immediately overcome even with mediation (e.g., teachers using exclusively organic scripts cannot avoid the aggravation of negative language transfer brought about by the interference of the IOS on SLA).

In order to discuss the phenomenon of the IOS, this work will explore some of the difficulties that emerge during the acquisition of the phonetic inventory of English. This is because it is presently humanity's most spoken language, with over 1.5 billion speakers around the world (Crystal 2008; Dyvik 2024), many of whom make use of it on a regular basis as a *lingua franca*. Likewise, due to its global importance and widespread presence, English is presently considered to be the language of science, the global economy, and pop culture (Chua 2022).

Although the thesis of IOS postulates that all naturally-evolving writing systems are unsuited for SLA, this work focuses on Latin-script alphabets for two reasons: their universal quality, since they build upon the foundation of the most utilized script in the world (Encyclopaedia Britannica 2024),

reaching at least 70% of the human population; and their being the group to which belongs the English alphabet, the script used to register the globally-important English language.

Furthermore, the study at hand focuses on some of the phonological hindrances caused by the IOS, since they belong to a subdomain that has not received the same attention as other areas of inquiry in SLA (Juffs 2011, 279).

With the aim of enabling teachers to engage in fruitful tasks that remain within their students' ZPD, the present study ultimately argues that the usage of organic scripts during SLA should necessarily be complemented by properly adapted concepts from linguistics, for "it is important for [language] instruction to be informed by research that links careful linguistic analysis" (Juffs 2011, 284).

Conceptualizing the Inability of Organic Scripts

On the convoluted circumventions of Latin-script alphabets

Alphabets based on the Latin script comprise the majority of the writing systems presently used in the world (Encyclopaedia Britannica 2024), which are prevalent in all of the Americas, the majority of Africa, Europe, Oceania, and in significantly populated parts of Asia, such as Indonesia, Malaysia, and the Philippines. The importance of the Latin script is further increased due to its being used as the basis for the alphabet of the widespread English language.

Having probably evolved from the Etruscan alphabet (Ullman 1927), the original purpose of the Latin alphabet was to register the local Latin language. It was only later, beginning with the expansion of the Roman Empire, that the script started disseminating across Europe, this being the reason why it came to register the majority of Romance languages, descendants of Vulgar Latin. Only centuries later, during the Age of Discovery, did the alphabet finally reach the Americas.

Since the Latin alphabet was not originally devised to register the Romance languages that only later came into existence (e.g., Italian, Spanish, Judeo-Spanish, Portuguese, French, Romanian, etc.), the usage of Latin-based scripts for this purpose can become significantly convoluted. This is comprehensible due to its ancient origins and surely acceptable in historical and sociocultural terms, but regarding second-language acquisition, the script's

limitations are often the cause of severe and unnecessary inefficiency.

To circumvent the difficulties that Latin-based alphabets face when used to register languages other than Latin, it often requires intensive use of digraphs, accents, diacritics, and special letters, among other adaptations; and in fact, even the currently standard Latin script holds graphemes that were not present in the original Latin alphabet, namely the letters *j*, *u*, and *w*.

As each Romance-speaking culture found their own ways to circumvent the limitations of the Latin script, profound disparities began to emerge between their languages, notwithstanding having a common linguistic ancestor. Among these circumventions, a select few are here cited to better illustrate the matter:

- the letter *ñ*, historically a ligature of the digraph *nn*, was added to the Spanish alphabet to represent the phoneme /ɲ/, which is in turn registered:
  - in Catalan via the digraph *ny*;
  - in French and Italian via *gn*;
  - in Occitan via *gn* and *nh*;
  - in Portuguese via *nh*;
- the letter *ș* is exclusive to the Romanian alphabet, where it represents the phoneme /ʃ/, which is in turn represented:
  - in Italian via the digraph *sc* when before *e* or *i*, and via the trigraph *sci* everywhere else;
  - in French via the digraph *ch*;
  - in Portuguese via either the digraph *ch* or the letter *x*;
- the letter *ț*, also exclusive to the Romanian alphabet, represents the phoneme /t͡s/, which is in turn registered in Italian as *z*;
- among Romance languages, the phoneme /ʎ/ is represented via:
  - the digraph *lh* in Portuguese;
  - the digraph *ll* in Spanish, which can also represent [j], [ʒ] or [dʒ], depending on the dialect;
  - the digraph *ll* in French (only applies for some dialects, as the phoneme represented by the digraph is commonly realized in the modern language as [j]).
- among Romance languages, the letter *x*, aside from representing its original sound of /k͡s/, came to register several different phonemes, among which a few are here cited:
  - /ʃ/, /s/, /z/, /k͡s/ or, more rarely, /g͡z/ in Portuguese;
  - /(k)s/, /ʃ/, /t͡ʃ/, /s/, or /x/ in Spanish;
  - /g͡z/, /k͡s/, /s/, /z/ in French, where it can also be silent;
  - /t͡s/, /s/, /g͡z/ or /g͡ʒ/ in Occitan.

Latin-script alphabets experience even more complex difficulties when used to register non-Romance languages. This is the case of English, whose complex phonetic inventory differs significantly in quality and quantity from those of various other languages, especially from those of Romance origin. However, as the English alphabet does not resort to any diacritics or special letters, the assortment of circumventions it requires to register the language it is intended to is considerably more extensive and complicated than those of the majority of other languages, among which a select few are here listed:

- vowel letters represent more than one actual vowel, e.g.:
  - *a* → /æ/, /ɑː/, /eɪ/, /ɛə/, etc.;
  - *e* → /ɛ/, /ɜː/, /iː/, /ɪə/, etc.;
  - *i* → /ɪ/, /aɪ/, /aɪə/, /ɜː/, etc.;
- vowel phonemes can be represented via different letters and letter combinations (digraphs, trigraphs, and tetragraphs), e.g.:
  - /iː/ → *i, e, ea, ee, ie, ey, eye,* etc.;
  - /uː/ → *oo, u, ue, oe, ew, ough*, etc.;
  - /ʌ/ → *u, o, e, oe, ou, a, au,* etc.;
  - /ə/ → *a, e, i, o, u, y, ae, ah, ei, eo, eou, oe,* etc.;
  - *oa* → /oʊ/, /oʊə/, ɔː/, /uːə/, etc.;
  - *oe* → /iː/, /oʊ/, /ɪ/, /uː/ , etc.;
  - *ui* → /ɪ/, /aɪ/, /uː/, /juːɪ/, etc.;
  - *ch* → /(t)ʃ/, /k/, /ʃ/ /tʃ/, etc.;
  - *ough* can represent over 10 different phonemes, some of which are demonstrated in the phrase "*though the tough cough and hiccough plough him through*". Its possible realizations include /ʌf/, /ɒf/, aʊ/, /oʊ/, /ɔː/, and /uː/.

Among the factors that have led to these complicated circumventions, one of the most relevant to this study is the abundance of consonants and vowels in the English language, which far outnumber the letters available in the Latin script. For instance, while the General American accent holds a total of 40 unique sounds (25 consonants, 12 vowels, and 3 diphthongs), the modern English alphabet counts with no more than 26 letters (19 for consonants, 5 for vowels, and 2 for semivowels). Moreover, as mentioned before, due to the absence of diacritics and special letters, the Latin-based English alphabet relies far more on orthography than the majority of languages that are registered with Latin-script alphabets, which has led to highly irregular spelling and complex orthography rules that can confuse even native speakers.

On the aggravation of negative language transfer by the IOS

The inherent incapacity of the Latin-based English alphabet to consistently represent the sounds of the English language makes the acquisition of the language's phonetic inventory considerably difficult, leading, among other issues, to the aggravation of negative language transfer.

The phenomenon of *language transfer* is a concept of great importance to SLA, with research on the topic going as far back as the 1950s (Zhang 2022, 38). It refers to the influences sprung from the similarities and differences between the target L2 and previously acquired languages, which affect all of the linguistic subsystems (e.g., semantics, syntax, phonology, phonetics) (Odlin 2003, 436-437).

In other words, during SLA, students are naturally prone to relying on systems from their L1 (first language) to interpret L2 input. To illustrate, learners might, for example, make use of the phonetic inventory of their L1 when attempting to speak or read in their target L2.

The L1 influences that help learners during SLA are referred to as *positive* (e.g., equivalent phonemes, similar phones, letters that represent the same phonemes), while those that create difficulties to learners are referred to as *negative* (e.g., discordant phonemes, dissimilar phones, letters that represent different phonemes).

Although language transfer is considered a normal phenomenon and an integral part of the development of the *interlanguage*—i.e., the idiolect that learners naturally produce during SLA, which contains features from both their L1 and their target L2—negative language transfer may, if uncorrected, ultimately lead to the internalization of incorrect foundations of the target L2, such as a defective phonetic inventory.

When this happens, learners may feel that their listening comprehension of their target L2 is insufficient or that the message they are trying to convey when speaking in that language does not get across. These factors have a significant possibility of leading to a diminished willingness to communicate and language anxiety, both of which work in a negative feedback loop to further hinder SLA and aggravate other existing issues.

Expanding on the idea, it has been established that language anxiety, an important psychosocial variable investigated by SLA researchers since the 1950s, can be aroused by the necessity of speaking and listening in a foreign

language while not being able to completely understand its words or speak without mistakes (Yamashiro and McLaughlin 2001, 113–115).

This sort of anxiety has manifold manifestations and is frequently negatively correlated to L2 performance, achievement, and willingness to communicate; in fact, anxious foreign language students often consider communicating in the L2 they are learning as among the most scaring moments in classrooms (Fujii 2021, 2-3; Fujii 2018, 48; Yashima 2002, 55).

In turn, a diminished willingness to communicate further increases language anxiety as students feel they are not able to properly handle the moments of communication in their target L2.

Considering that language transfer essentially originates from language learners' attempts to rely on L1 systems to interpret L2 input, it is reasonable to assume that organic scripts will always, in a way or another, mislead learners. For example, while the Latin letters are mostly the same across languages, they may actually represent completely different phonemes.

To illustrate this, let us consider the Latin letter *t*, which represents voiceless dental and alveolar plosives [t] and [t̪] across many languages. The /t/ phoneme in English, however, is aspirated as [tʰ] in several instances, a realization that is not represented in any way by the Latin letter. This means that native speakers of Spanish, for example, would tend to read the word *top* as [tɔp] or [top] instead of the correct [tʰɑp]; and as it can be noted, even the vowels represented by the same letter *o* differ substantially.

A more remarkable example is the case of the letter *r*, which generally represents either voiced alveolar taps [ɾ] and trills [r] or voiced uvular fricatives [ʁ] and trills [ʀ]: in English, the letter registers the voiced postalveolar approximant [ɹ̠], a sound found in far fewer languages, which is additionally often labialized as [ɹ̠ʷ]. This means that the word *red*, pronounced as [ɹ̠ʷɛd] in English, could be read by a speaker of Spanish as [rɛd]; a speaker of Brazilian Portuguese, however, would presumably read it as [ˈʁɛ.di]—or even [ˈʁɛ.d͡ʒi], depending on the regional accent—since the native speakers of said language tend to color consonants in the position of syllabic coda with an [i] sound.

As the Latin alphabet, *per se*, does little to consistently and accurately guide learners on how to pronounce words of any language, it rests predominantly on the shoulders of language teachers to mentor students in memorizing complex

orthography rules and different sounds for Latin letters they already know. For example, teachers will have to instruct students that the letter *t* in English is aspirated in several instances or that the English *r* registers a considerably rare phoneme that is likely non-existing in their L1 phonetic inventory.

This approach is problematic because, besides requiring language students to memorize different sounds for letters they already know, which can be difficult for novice L2 learners, it provides minimal support in helping students understand how to produce the sounds that are unfamiliar to them.

These are among the reasons why this work deems IOS to involve some of the "many aspects about language transfer in second language acquisition waiting to be researched" (Zhang 2022, 41).

On the Elaboration of Linguistically Expanded Scripts

The present work argues that the lack of usage of tools and concepts from linguistics by language teachers and instructors is among the major reasons why language learners remain highly vulnerable to the IOS.

For this reason, this work advocates for the pedagogical adaptation of concepts from linguistics, which can serve as substrata for the elaboration of teaching strategies that are capable of bypassing the IOS during SLA, thus bringing the foundations of the target L2 back to learners' ZPD during crucial stages of its acquisition process.

It is already established that students can assimilate concepts from a variety of scientific fields when properly adapted (e.g., photosynthesis, atomic theory, chemical reactions). Language teaching, however, to this day lacks adequate scientific support and generally does not make considerable use of concepts and tools offered by the field of linguistics, especially in classrooms.

Seeking to contribute to increasing scientific linguistic input in language teaching, the present work advocates for the use of an adaptation of the International Phonetic Alphabet (IPA), here referred to as Linguistically Expanded Scripts, during SLA.

Created by the International Phonetic Association, founded in Paris in 1886 by language teachers who advocated for a "phonetic notation to be used in schools as a method of helping children to acquire a realistic pronunciation of foreign languages" (International

Phonetic Association 1999, 194), the IPA stands as an unequivocally scientific writing system, specifically designed to register the vast array of sounds found in human languages.

In order to illustrate how the IPA allows for a clear rendering of the words that otherwise suffer from the low sound-to-grapheme ratio in the English alphabet, the forementioned English phrase used to demonstrate the different pronunciations of the *ough* tetragraph is presented below alongside a transcription to the IPA:

ENG	though the tough cough and hiccough plough him through
IPA	ðoʊ ðə tʌf kɑf ænd ˈhɪ.kʌp plaʊ hɪm θɻuː

As the need to learn every one of the IPA symbols would bring sizable and unproductive difficulty to language learners, the present work advocates for the elaboration of Linguistically Expanded Scripts (LES) to be used during SLA in order to help overcome the obstacles brought about by the IOS.

LES are to be reduced versions of the IPA; that is, alphabets containing all the letters of the standard Latin script, plus the IPA graphemes considered to be strictly necessary to provide an adequate description of the phonetic inventory of the target L2. As the letters to be added are reasonable in number, the task of memorizing a LES is expected to remain within learners' ZPD.

Exemplifying, the LES for the acquisition of General American English would require a total of 37 letters (23 for consonants and 13 for vowels). Out of these, only 15 graphemes (namely, ŋ, ʤ, ð, ʃ, ʒ, ɪ, ɹ, ʍ, ʊ, ə, æ, ʌ, ɛ, ɔ, and ɑ) would present themselves as new letters for students that are already familiar with the Latin script; and 11 among these 15 are simple variants of Latin letters, which appear either rotated or only slightly modified, as shown below:

- ɑ, æ → *a*;
- ð → *d*;
- ə → *e*;
- ɪ → *i*;
- ŋ → *n*;
- ɔ → *o*;
- ɹ → *r*;
- ʃ → *f* (a "long s");
- ʌ → *v*;
- ʍ → *w*;
- ʒ → *z*.

Along with these new letters, the inclusion of a few suprasegmentals and diacritics of easy memorization would further increase the efficiency of the LES, among which the following are by this work considered essential:

- ˈ, *primary stress marker*; is similar to an apostrophe (‘);

- ː, *long vowel and gemination marker*; is similar to a colon (:);
- ʰ, *aspiration marker*; a miniature *h*;
- ʷ, *labialization marker*; a miniature *w*.

From the abovementioned inclusions to the Latin alphabet results the Linguistically Expanded General American Script (LEGAS). The number of possible LES is expected to be at least equal to that of existing languages.

All that said, the present work advocates that language teachers should employ learners' target L2 LES at least in the following ways during SLA:

- when presenting new vocabulary;
- when instructing and correcting pronunciation;
- by devising exercises where learners need to transcribe sentences to and from their target L2 LES.

Conclusion

The present work sought to describe the concept of the Inability of Organic Scripts, which refers to the inherent incapacity of naturally-evolving writing systems to adequately mediate SLA. Limiting the scope to the Latin script, which the majority of writing systems worldwide use as their basis, this work discussed how the IOS pushes important foundations of language learners' target L2 out of their ZPD during a pivotal stage of SLA, ultimately contributing to generate or aggravate several issues, among them negative language transfer and language anxiety, which were here briefly discussed.

In order to bypass the detrimental effects caused by the IOS during SLA, this work advocates for the use of Linguistically Expanded Scripts, scientifically-oriented writing systems expected to work, in a sense, as abridged versions of the IPA.

LES are produced by supplementing the standard Latin script with the IPA graphemes strictly necessary for an efficient acquisition of the phonetic inventory of the target L2, i.e., the symbols from the IPA that represent the phonemes from the target L2 that do not have a corresponding letter in the standard Latin script.

It is of importance to note that although the work at hand, for the sake of brevity, focused on some of the phonological difficulties caused by the IOS during SLA, paramount to the thesis here postulated is the understanding that the IOS either causes or aggravates issues on other linguistic subsystems as well,

which are to be discussed by future studies on the topic.

Acknowledgments

I would like to extend my sincere gratitude to Marta Sanches and Eliana Storgatto, whose generous patronage has made this research possible.

References

Chua, Amy. 2022. "How the English Language Conquered the World." The New York Times. Last modified January 18, 20. https://www.nytimes.com/2022/01/18/books/review/the-rise-of-english-rosemary-salomone.html

Crystal, David. 2008. "Two Thousand Million?" *English Today* 24, no. 1 (March): 3–6. https://doi.org/10.1017/S0266078408000023.

Dyvik, Einar. 2024. "The Most Spoken Languages Worldwide in 2023 (By Speakers in Millions)." Statista. Last modified March 22, 2024. https://www.statista.com/statistics/266808/the-most-spoken-languages-worldwide/

Encyclopaedia Britannica, The Editors of. "Latin alphabet." Encyclopaedia Britannica. Last modified February 20, 2024. https://www.britannica.com/topic/Latin-alphabet

Gass, Susan, and Larry Selinker. 2001. *Second-Language Acquisition: An Introductory Course*, 2nd ed. Mawah: Lawrence Erlbaum Associates.

International Phonetic Association. 1999. *Handbook of the International Phonetic Association: A Guide to the Use of the International Phonetic Alphabet*. Cambridge: Cambridge University Press.

Juffs, Alan. 2011. "Second Language Acquisition." *WIREs Cognitive Science* 2, no. 3 (April): 277–86. https://doi.org/10.1002/wcs.106

Kurt, Serhat. 2020. "Vygotsky's Zone of Proximal Development and Scaffolding." Educational Technology. Last modified August 18, 2020. https://educationaltechnology.net/vygotskys-zone-of-proximal-development-and-scaffolding/

Odlin, Terence. 2003. "Cross-Linguistic Influence." In *The Handbook of Second Language Acquisition*, edited by Catherine Doughty and Michael Long, 435–86. Malden: Blackwell Publishing.

Ullman, Berthold. 1927. "The Etruscan Origin of the Roman Alphabet and

the Names of the Letters." *Classical Philology* 22, no. 4 (October): 372–377. https://www.jstor.org/stable/262978

VanPatten, Bill, and Alessandro Benati. 2010. *Key Terms in Second Language Acquisition.* London: Continuum International Publishing Group.

Yamashiro, Amy, and John McLaughlin. 2001. "Relationships among Attitudes, Motivation, Anxiety, and English Language Proficiency in Japanese College Students". In *Second-Language Acquisition Research in Japan*, edited by Peter Robinson, Mark Sawyer and Steven Ross, 113-27. Tokyo: Japan Association for Language Teaching. https://www.researchgate.net/publication/259868644

Yashima, Tomoko. 2002. "Willingness to Communicate in a Second Language: The Japanese EFL Context." *Modern Language Journal* 86, no. 1 (December): 54–66. https://doi.org/10.1111/1540-4781.00136