LATEL: a Logical And Transparent Experimental Language

Kelvin M. Liu-Huang
Carnegie Mellon University
ecesare@gmail.com

1. Introduction

Previous constructed languages were built on etymological (Zamenhof, 1887), semantic (Bliss, 1965), precision (Lojban), or other academic merit. Few have addressed cognitive benefits. The arbitrary etymology of most natural languages creates cognitive dissonance. Consider how mathematical expressions can precisely express a great deal while using a very small number of definitions. Compared to the low precision and extent of knowledge required to learn natural languages, mathematical expressions seem better in these ways. The tradeoff is, of course, that mathematical descriptions can be very confusing or unwieldy. We attempt to address all these concerns in order to construct a phonosemantographic spoken language. Language should ideally harmonize speech, listening, reading, writing, and comprehension in order to facilitate learning. By infusing each letter with meaning and using phonetic orthography, sound, spelling, and meaning can all be inferred from each other, reducing ambiguity, speeding up learning, and even allowing efficient and deterministic creation of neologisms. The orthography is nearly identical to the IPA symbols for the phonemes themselves for simplicity. Consonants represent the set of all objects of a certain type. Vowels represent Boolean logic, set logic, and arithmetic operator (often all at once because many operations have analogous operations for Boolean/set/numerical types). Words are formed by selecting subsets containing the desired objects. Morphology must be inorder because it is sometimes impossible to pronounce preorder and postorder due to consonant duplicates (vowel clusters also pose a problem).

2. Phonetics

Phonemes for LATEL were greedily selected in order of the most prevalent phonemes among languages worldwide (Moran & McCloy). However, some were discarded due to similarity with previously selected phonemes.

Most of the selected vowels, i, y, ɯ, u, e, o, ɛ, ɔ, a, and ɑ (indicated hereafter using IPA), coincide with the IPA vowel gridlines, which benefit from high sound contrast. The other selected vowels, ı, ʊ, and ə, are also quite phonetically and spatially distinct. Additionally, operations which have a tendency to neighbor other vowels were assigned vowels which correspond to semivowels.

The voiceless consonants, k, p, s, t, h, f, ʃ, and θ, were all selected before their voiced counterparts, g, b, z, d, ɦ, v, ʒ, and ð, due to ease of articulation. No approximants were selected because these are easily confused with vowels.

Table 1. All LATEL letters, their phonemes (indicated with IPA), and their definitions.

LATEL	IPA	Boolean algebra	set algebra	set algebra output type	scalar algebra
i	i		@ $\stackrel{\text{def}}{=}$ set in most recent ancestral :, ∃, or ∀	set	
			@P $\stackrel{\text{def}}{=}$ previous set of class P	set	
			P@Q $\stackrel{\text{def}}{=}$ Qth previous set of class P	set	
a	a		∃P $\stackrel{\text{def}}{=} \exists p \in \$(P)$	Boolean	
			P∃QR… $\stackrel{\text{def}}{=} \exists p \in \$(P)(Q, R, \ldots)$	Boolean	
u	u		. $\stackrel{\text{def}}{=} i(@, @)$	number	
			.Q $\stackrel{\text{def}}{=} \{p \mid p \in @, i(@, @) = Q\}$	set	
			P.Q $\stackrel{\text{def}}{=} \{p \mid p \in P, i(p, P) = Q\}$	set	
			.P $\stackrel{\text{def}}{=} i(P, P)$	number	
o	o	P∨Q	P∪Q	set	P+Q
e	e	P∧Q	P∩Q	set	P×Q
ɛ	ɛ	P→Q	P⊂Q	Boolean	P<Q
ɔ	ɔ	P←Q	P⊃Q	Boolean	P>Q
ə	ə		P:QR… $\stackrel{\text{def}}{=} \cup \{p \mid p \in \$(P), Q, R, \ldots\}$	set	
ɪ	ɪ	P↔Q	P=Q	Boolean	P=Q
ʊ	ʊ	P⊕Q	PΔQ	set	P≠Q
y	y	¬P	^{C}P	set	-P
w	ɯ		#P $\stackrel{\text{def}}{=} \lvert P \rvert$	number	
α	ɑ		P∀QR… $\stackrel{\text{def}}{=} \forall p \in \$(P)(Q, R, \ldots)$	set	

where $\$(P) = \{\{p \mid p \in P, i(p, P) = j\} \mid j \in I(A)\}$, $I(A)$ is the set of A-indices, and $i(a, A)$ is the A-index of a

LATEL	IPA	definition	recommended set	recommended indices	0th element
m	m	mass	all particles of mass-energy	particles	
k	k	concepts	all knowledge in the mind and concepts	concepts	
p	p	position	all particles and entities	Planck volumes	here
n	n	enthalpy	all particles of mass-energy	particles	
s	s	organisms	all particles belonging to living entities	entities	I
t	t	time	all particles and entities	each instant	now
b	b	(unassigned)			
h	h	souls/minds	all souls/minds	souls/minds	my mind
g	g	(unassigned)			
ŋ	ŋ	(unassigned)			
r	r	entropy	all particles of mass-energy	particles	
z	z	integers*			
v	v	things	all particles belonging to non-living objects	entities	
ɣ	ɤ	bodies	all physical particles	entities	my body

3. Orthography

The orthography of LATEL is phonemic (each written letter corresponds to a single phoneme and vice-versa). As shown in Table 1, most of the letters in LATEL simply correspond to the IPA symbols for their phonemes. However, a few letters do not correspond to IPA for the sake of legibility or simplicity.

4. Morphosyntax

Like in most languages and mathematical expressions, LATEL contains objects and relationships, can represent the order of those operations/relationships in a tree, and produces statements which the speaker claims is true. In LATEL, each consonant represents a set of elementary objects, while each vowel represents a Boolean, set, or scalar operation. By applying operations to consonants, it's possible to create a variety of new sets. Each expression in LATEL is necessarily a statement that the speaker purports as true. Upon forming an expression by applying operations to various objects, the speaker can express various beliefs.

Like any other language, since the structure is tree-shaped, we must serialize the tree to make it possible to dictate. Like English and many other languages, we use inorder traversal of the tree to convert the tree into a sequence. Inorder traversal is typically only defined for binary trees. To generalize to arbitrary trees, we define that half of the children (rounded down) of each branch are on the left, and the rest are on the right. Figure 1 shows an example LATEL expression tree which is then transcribed through inorder traversal into a sequence.

Because information is lost during inorder serialization, we cannot deterministically deduce the original tree from the spoken sequence. Hopefully, the tree structure can be deduced through context. However, if it cannot, then the structure can be clarified using pitches; the speaker can sing the expression by selecting a pitch for each vowel that is lower than the pitch for the vowels in higher branches.

Preorder or postorder traversal preserve enough information during serialization such that the tree structure can be recovered. However preorder traversal has a tendency to create consonant clusters which are difficult to articulate; in particular, consonant pairs are completely impossible to enunciate clearly. Vowel clusters are also common and more difficult to articulate. Postorder traversal suffers from the same problem. Additionally, the vowels (operators) appear at the end of the sequence, so the listener does not know which consonants (objects) belong to which operations until the end of the sentence, which can be very confusing. Also, both preorder and postorder serialization are rarely used among natural languages, which would

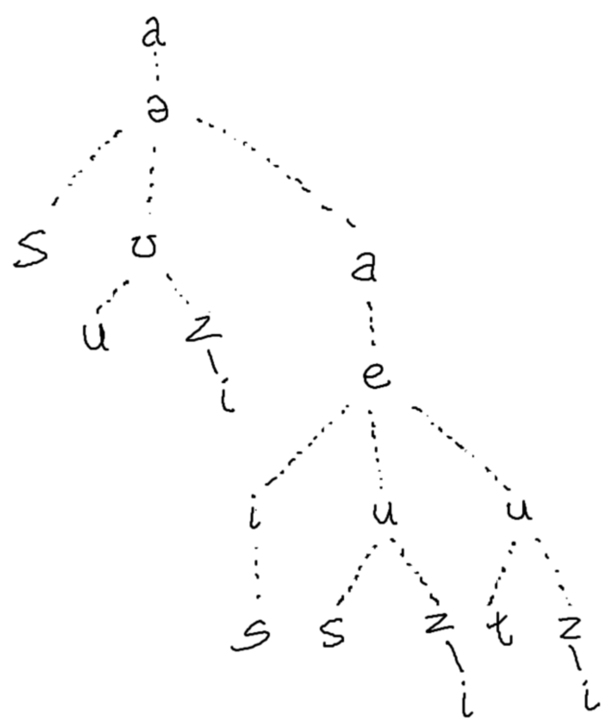

Figure 1. Example expression in LATEL. The expression transcribes into "s ə u ʊ zi a i p e s u zi t u zi," which literally translates to "there-exists-at-least-one ({ organisms such that ((they) is-not (me)) and (there-exists-something-in (the intersection of (the (organism)), (the (organism) who is (zero)th), and (the (time) which is (zero)th))) })," roughly translating to "I am accompanied".

also impose a barrier to learning.

We chose not to include any grammatical inflection to keep the linguistic rules simple.

Note that a formal distinction between morphology and syntax does not exist in LATEL because individual letters already bear meaning. Instead, speakers are encouraged to form words out of convenience. For example, in Figure 1, one of the subtrees is "s u zi," literally translating to "the (organism) who is (zero)th," which is defined as the speaker, thus roughly translating to "I" (see Semantics). Thus it would be reasonable to define "I" as "suzi." Words can then be delimited by spaces in writing and pauses in speech.

5. Semantics

5.1. Sets (Consonants)

Each consonant in LATEL corresponds to the set of all elements in the set of all objects of a certain type (see Table 1). However, the elements and nature of these objects may vary depending on the beliefs of the speaker. For example, "m," defined as mass, might represent the set of all elementary particles in the universe, or perhaps the set of all strings. "s," defined as organisms, might represent the set of all particles belonging to living entities (indexed by individual entities). "t," defined as time, might represent the set of everything (index by timed). Performing operations on these elementary sets lets us build more complex sets. For example, taking the intersection of the 0^{th} index of organisms (I) and the 0^{th} index of time (now) gives us just set of objects inside the person I am now.

5.2. Set Indexing

In order to build sets with appropriate subsets, sets are indexed depending on the set class (see the column in Table 1). For example, even though time and space both contain everything, the objects in time are indexed by moments, while the objects in space are indexed by position. To enable this scheme mathematically, all set-building operations, :, ∃, and ∀, use the specialized definition, $\$(P) = \{\{p|p \in P, i(p,P) = j\}\big|j \in I(A)\}$, where $I(A)$ is the set of *A*-indices and $i(a,A)$ is the *A*-index of *a*.

Because many of these sets are infinite, we needed to define default values in order to meaningfully select elements from sets and express useful statements. Thus for many sets, the 0^{th} element is specifically defined (see the column in Table 1).

5.3. Operations (Vowels)

Each vowels encodes a logical, set, and/or scalar operation (see Table 1). These behave as expected. For example, $P \wedge Q$ (which would be transcribed as "peq" in LATEL) is boolean variable which is true if both *P* and *Q* are true. *P*.*Q* is the subset of elements in *P* which all have *P*-index equal to *Q*.

Since many logical operations have set and scalar algebraic analogs, we overloaded these operators. If the children of these operators are Booleans, sets, or scalars, then the operator applies the corresponding type of operation. Usually the output type is the same as the input, but for some set algebra operations, the output is not a set; instead they are indicated in the "set algebra output type" column.

Examples

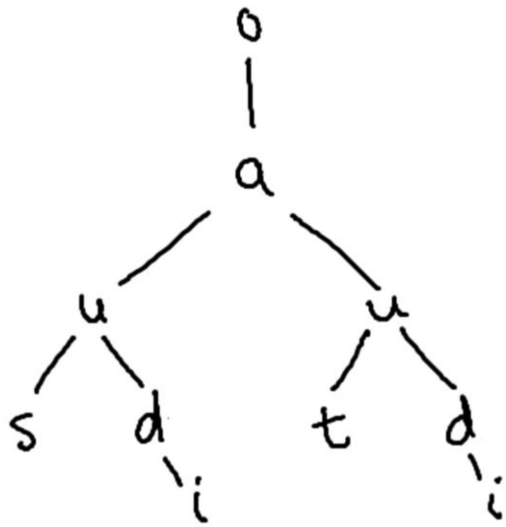

Figure 2. Example expression in LATEL. The expression transcribes into "o sudi a tudi," which literally translates to "there-exists-something-in (the intersection of (the (organism) who is (zero)th) and (the (time) which is (zero)th)))," roughly translating to "I exist at present."

In LATEL, deep philosophical concepts such as in Figure 2 are very easy to express. One disadvantage is that mundane everyday concepts which are of no interest anyways, such as in Figure 3, are difficult to express.

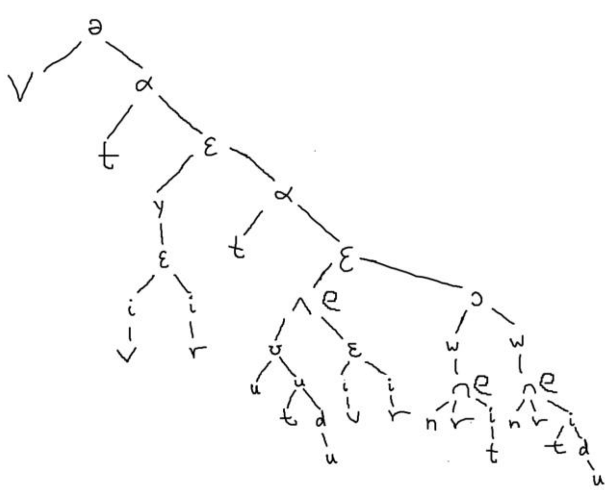

Figure 3. Example expression in LATEL. The expression transcribes into "vətαyivɛirɛtαuʊtudueivɛirɛwneritɔwnertidu," roughly translating to "food." The literal translation is left as an exercise to the reader.

Acknowledgements

The author would like to thank Cory Stevenson for numerous discussions, Junxing Wang for insight and discussions on complexity theory and data structures, and the academy.

Bibliography

Bliss, C. K. (1965). *Semantography, a non-alphabetical symbol writing, readable in all languages; a practical tool for general international communication, especially in science, industry, commerce, traffic, etc., and for semantical education, based on the principles of ideog.* Sydney: Institute for Semantography.

Moran, S., & McCloy, D. (n.d.). *Segments*. Retrieved 2019, from Phoible 2.0: phoible.org/parameters

Zamenhof, L. L. (1887). *Dr. Esperanto's International Language: Introduction & Complete Grammar.*