LATEL: a Logical And Transparent Experimental Language

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

1. Introduction

Most previous constructed languages intended for human use set out to improve etymological integrity (Zamenhof, 1887), semantic clarity (Bliss, 1965), consistency (Weilgart, 1979; Cowan, 1997; Quijada, 2004), or other academic merits. Not many (Weilgart, 1979; Cowan, 1997; Bourland & Johnston, 1991; Quijada, 2004; Lang, 2014) have addressed cognitive benefits. First, the arbitrary phonetics and morphology of most natural languages creates cognitive dissonance, which can be easily averted. Also consider how mathematical expressions can precisely express a great deal using a very small number of definitions. Compared to the ambiguity and learning barrier of natural languages, mathematical expressions seem better in these ways. The tradeoff is, of course, that mathematical descriptions can be very elaborate or unwieldy.

We attempt to address all these concerns in order to construct LATEL, a phonosemantographic spoken language. Language should ideally harmonize speech, listening, reading, writing, and comprehension in order to facilitate learning. Like aUI (Weilgart, 1979), by infusing individual letter with meaning and using phonetic orthography, letters, sounds, and meaning can all be inferred from each other, reducing ambiguity, speeding up learning, and even allowing efficient and deterministic creation of neologisms. For simplicity, the orthography mostly matches the IPA symbols for the phonemes themselves. Unlike aUI (Weilgart, 1979), LATEL attempts to express semantics entirely through logic rather than metaphor. Semantically, consonants represent the set of all objects of a certain class. Vowels represent Boolean, set, and/or scalar algebraic operators (possibly multiple all at once because many operations have analogous operations for Boolean/set/scalar types, and the operator is overloaded). Expressions are formed by selecting subsets containing the desired objects. Morphology must derived from inorder traversal because it is impossible to pronounce preorder and postorder of many trees due to consonant duplicates (vowel clusters also pose a problem).

Not only do we shape language, the Sapir-Whorf hypothesis suggests that language also influences (or perhaps determines) our thoughts and behavior. Some previous conlangs (Weilgart, 1979; Cowan, 1997; Bourland & Johnston, 1991; Quijada, 2004) have attempted to explore or utilize this hypothesis to improve cognitive function, but most achieve this by through increased complexity (Quijada, 2004). Meanwhile, Lojban is largely grounded in logic, though word formation is still arbitrary because they are synthesized from existing languages (Cowan, 1997). LATEL attempts to ground both morphology and syntax in pure logic using the same building blocks as propositional logic, set theory, and arithmetic.

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

LATEL	IPA	Boolean algebra	set algebra	scalar algebra
i	i		$@ \stackrel{\text{def}}{=}$ current set (p in most recent ancestral :, ∃, or ∀) $@P \stackrel{\text{def}}{=}$ previous set of class P $P@Q \stackrel{\text{def}}{=}$ Qth previous (inorder) set of class P	
a	a		$\exists P \stackrel{\text{def}}{=} \exists p \in Z(P)$ $P\exists QR\ldots \stackrel{\text{def}}{=} \exists p \in Z(P)(Q, R, \ldots)$	
u	u		$. \stackrel{\text{def}}{=} j(@)$ $.Q \stackrel{\text{def}}{=} \{p \mid p \in @, j(@) = Q\}$ $P.Q \stackrel{\text{def}}{=} \{p \mid p \in P, j(p) = Q\}$ $.P \stackrel{\text{def}}{=} j(P)$	
o	o	$P \vee Q$	$P \cup Q$	$P+Q$
e	e	$P \wedge Q$	$P \cap Q$	$P \times Q$
ɛ	ɛ	$P \rightarrow Q$	$P \subset Q$	$P<Q$
ɔ	ɔ	$P \leftarrow Q$	$P \supset Q$	$P>Q$
ə	ə		$P{:}QR\ldots \stackrel{\text{def}}{=} \cup \{p \mid p \in \$(P), Q, R, \ldots\}$	
ɪ	ɪ	$P \leftrightarrow Q$	$P=Q$	$P=Q$
ʊ	ʊ	$P \oplus Q$	$P \Delta Q$	$P \neq Q$
y	y	$\neg P$	${}^{C}P$	$-P$
w	ɯ		$\#P \stackrel{\text{def}}{=} \lvert P \rvert$	
α	ɑ		$P\forall QR\ldots \stackrel{\text{def}}{=} \forall p \in Z(P)(Q, R, \ldots)$	

where $Z(P) = \{\{p \mid p \in P, j(p) = i\} \mid i \in J(A)\}$, $J(A) \subseteq \mathbb{Z}$ is the enumeration of *A* using the recommended indices, and $j: J(A) \rightarrow A$ gives the element in *A* as enumerated by $J(A)$

LATEL	IPA	class	recommended set	recommended indexing, $J(A)$	0th element
m	m	mass	all particles of mass-energy	particles	
k	k	concept	all knowledge in the mind and concepts	concepts	
p	p	position	all objects	Planck volumes	here
n	n	enthalpy	all particles of mass-energy	particles	
s	s	organism	all objects belonging to living entities	entities	I
t	t	time	all objects	Planck times	now
b	b	unassigned			
h	h	soul/mind	all souls/minds	souls/minds	my mind
g	g	unassigned			
ŋ	ŋ	unassigned			
r	r	entropy	all particles of mass-energy	particles	
z	z	integer*			
v	v	thing	all objects belonging to non-living objects	entities	
ɣ	ɣ	body	all particles of mass-energy	entities	my body

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.

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. Morphology and Syntax

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,

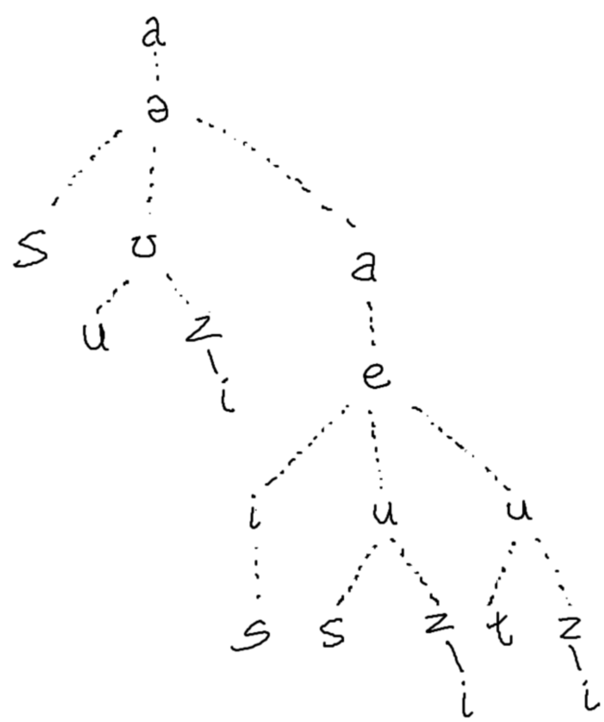

Figure 1. Example expression in LATEL. The expression transcribes into "s ə u ʊ zi a i p e s u zi t u zi," which translates literally to "there-exists-at-least-one ({ organisms such that ((they) is-not (me)) and (there-exists-something-in (the intersection of (the (organism)), (the (organism) who is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

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.

Alternatively, preorder or postorder traversal would 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, preorder serialization is almost never and postorder serialization is rarely found among natural languages, which could 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 each letter is already a complete sememe. Instead, speakers are encouraged to form words merely for convenience. For example, in Figure 1, one of the subtrees is "s u zi," literally translating to "the (organism) who is (zero)th," which corresponds to "I" (see Semantics). Thus it would be reasonable to define "I" as "suzi." Words are 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 class (see Table 1). However, the elements and nature of these objects may vary depending on the beliefs of the speaker. We do not claim to know all the answers to metaphysics. For example, "m," defined as mass, might represent the set of all elementary particles in the universe, or perhaps the set of all strings. A speaker may choose not to use "h," defined as souls/minds, if the speaker does not believe those exist. "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, $Z(P) = \left\{\{p|p \in P, j(p) = i\}\middle| i \in J(A)\right\}$, where $J(A) \subseteq \mathbb{Z}$ is the enumeration of *A* using the recommended indices and $j: J(A) \to A$ gives the element in *A* as enumerated by $J(A)$.

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 vowel encodes a logical, set, and/or scalar operation (see Table 1). These behave as expected. For example, *P*∧*Q* (which would be transcribed as "peq" in LATEL) is a Boolean expression which is true if both P and Q are true. $P.Q$ is the subset $X \subseteq P: \forall x \in X(j(x) = Q)$.

Since many logical operations have set and scalar algebraic analogs, we overloaded these operators. The operator applies the corresponding operation depending on whether its children are Booleans, sets, or scalars. Usually, but not always, the output type is the same as the input.

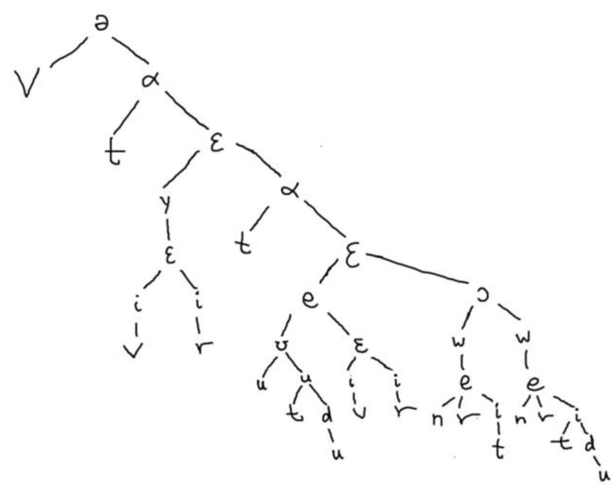

Figure 2. Example expression in LATEL, which transcribes into “vətɑyivɛirɛtɑuʊtudueivɛirɛwneritɔwnertidu,” roughly translating to “food.” The literal translation is left as an exercise to the reader.

6. Examples

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

Acknowledgements

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

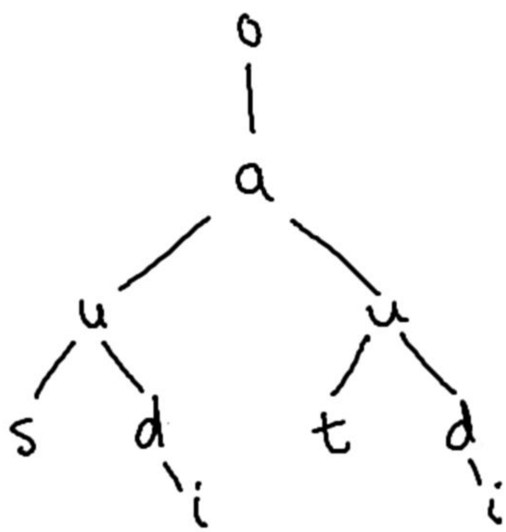

Figure 3. Example expression in LATEL, which transcribes into “o sudi a tudi,” which translates literally to “there-exists-something-in (the intersection of (the (organism) who is (zero)th) and (the (time) which is (zero)th)))” or roughly to “I exist at present.”

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.

Bourland, D. D., & Johnston, P. D. (1991). *To be or not: An E-prime anthology.* Institute of GS.

Cowan, J. W. (1997). *The complete Lojban language* (Vol. 15). Logical Language Group.

Lang, S. (2014). *Toki Pona: The language of good.* Tawhid.

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

Quijada, J. (2004). Retrieved 2019, from Ithkuil: A Philosophical Design for a Hypothetical Language: http://www.ithkuil.net/

Weilgart, W. J. (1979). *aUI: the language of space, Pentecostal logos of love & peace: for the first time represented and adapted to the needs of this planet.* Cosmic Communication Co.

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