L2 Phonology at the Interfaces

John Archibald

International Symposium on Monolingual and Bilingual Speech

2017
SLA Theory

Facing 3 epistemological challenges
Plato’s Problem

Knowing things that are not found in the environment

- Poverty of the Stimulus
- Codas
- Moras
- Traces
- Extrametricality
Orwell’s Problem

*Not* knowing things that are frequent in the environment

- Evidence for Evolution
- Evidence for Climate change
- $[\theta]$
Escher’s Problem

The challenge of Augmented Reality

- Perceiving things that aren’t in the acoustic (or visual) input
- Illusory vowels
Our Goal: A Unified Model of SLA

- Domain-specific hierarchical representations which are consistent across interfaces
- And a processor to drive performance and learning

See Archibald (2017a) for broader discussion
Representational (aka Indirect) Realism

- Essentialism
  - versus Externalism (Bresnan & Ford, 2010);
  - or Emergentism (MacWhinney & O’Grady, 2015)

- Phonology as cognition (Hale & Reiss, 2000)
Relevant Interfaces

- Phonetics
- Morphology
- Syntax
Rampant Heterogeneity: Mix and Match Theory

Understandably, each domain can have its own rich literature and vocabulary.
“If the duration of the closure extends beyond 50 ms, the intraoral pressure reaches a steady value equal to the subglottal pressure, and the glottal airflow decreases to essentially zero.”

--Stevens, (2000: 328)
Phonology

“This stochastic OT grammar....will be translated into an ordinal OT ranking...by randomly choosing a one-time value for each constraint from under the probability curve.”

--Tessier, (2016: 370)
Morphology

“There is a well-known hierarchy of grammatical persons in the [Algonquian] languages that determines the direct and inverse forms of the transitive verb. We may represent this hierarchy as follows: 2 > 1 > 3 > 3', which means that second person takes precedence over first, and these two take precedence over third proximate, which in turn takes precedence over third obviative”

--Dawe-Sheppard and Hewson, (1990:1)
Syntax

“We argue that impoverished versions of T and ν in VSO clauses lack the probe features involved in subject agreement, EPP, object shift and nominative/accusative valuation with Xhosa SVO sentences.”

--Carstens & Mletshe (2015)
But what are the commonalities?

What if we could bring it all together…
A single engine

Canada’s OTTER
Noted Interface Phenomena

- Variability in production
- Indeterminate knowledge
- Bottleneck of morphosyntactic accuracy (despite acquisition of core (‘narrow’) syntax and semantics)
Interfaces Reveal Architecture

- Single grammatical engine for phonology, morphology, and syntax
- Consistent with Bottleneck Hypothesis (Slabakova, 2014) – extended to ‘Narrow’ Phonology
Narrow Phonology

There is much evidence to show that L2 phonological categories can be acquired.

- Phonemes
- Syllable structure
- Vowel harmony
- Stress
The Unifying Theme Today

- The import and centrality of L2 phonology to SLA theory; part of GenSLA
- Successful L2 spell-out of categorical, representational phonology at key modular interfaces
The Centrality of Phonology

Now, you’d think talking about the importance of phonology at an international speech conference would be like
Coals to Newcastle
Architecture of an Interface

- Phonology/Morphology
- Phonology/Phonetics
- Phonology/Syntax
“The idea that a large part of the phonological grammar operates in ways that are utterly indifferent to or incompatible with the system for generating complex objects is suspicious, or at the least unfortunate… …rather than assuming that morphosyntax and morphophonology might be fundamentally different….it should be assumed that there is no extreme difference between these facets of grammar.”

--Embick (2010)
Properties of Distributed Morphology (DM)

- Functional morphemes are bundles of features (e.g., [past]) in the syntax which, via Vocabulary Insertion (VI), are spelled out phonologically.

- There is competition for allomorph selection but, crucially, no competition between complex objects (as in OT).
A syntactic derivation is sent to Spell Out which is then sent to both PF (Phonetic Form) and LF (Logical Form)
Properties of Distributed Morphology (DM)

There is a matrix of features on the syntactic terminal node and various Vocabulary Items would compete for insertion by seeing which affix matched the most features.
Properties of Distributed Morphology (DM)

Roots

- The store (of category-neutral roots) contains no phonological information (reminiscent of lemmas which have conceptual structure but no phonological structure)
- √dog or √chien or √σκύλος
- Grammatical categories are established in the syntax via functional heads such as ν or n.
DM & Language Mixing

Alexiadou et al. (2015) look at heritage Norwegian speakers in the U.S. who can mix English roots with Norwegian affixes

- Så play-de dom game-r
then play-PAST they game-INDEF.PL
Then, they played games.
DM & Language Mixing

The affixes come from the terminal nodes on a Norwegian syntactic structure (note V2)

Note, though, that an L2 root can get inserted into an L1 syntactic skeleton

Archibald (2016) suggested that roots can be subject to the same competition in bilinguals.
**DM Redux**

Consistent with Libben’s (2000) Homogeneity Hypothesis, the DM lexicon is non-selective.

See also Hilderman (2017) for an instantiation of DM in Sharwood Smith’s MOGUL to account for intra-word codeswitching.
Competition for Root Insertion

Haugen & Siddiqi (2013) argue that there is competition for roots and therefore the root is part of the Vocabulary list (see also Marantz (1995) on late Root insertion).

We also see evidence for this in the psycholinguistics of bilingual root insertion.
## Bilingual Competition

Studies on the non-selective bilingual lexicon (e.g. Dijkstra et al, 1999)

<table>
<thead>
<tr>
<th>Interlingual Homophones</th>
<th>Interlingual Homographs</th>
</tr>
</thead>
<tbody>
<tr>
<td>e.g. English/Dutch</td>
<td>e.g. English/Dutch</td>
</tr>
<tr>
<td>[lif] ‘leaf’/ ‘dear’</td>
<td>“glad” [glæd]/[xlat]</td>
</tr>
<tr>
<td>-slower (inhibited)</td>
<td>-faster activation</td>
</tr>
<tr>
<td>activation</td>
<td></td>
</tr>
<tr>
<td>-don’t share a root</td>
<td>-don’t share a root</td>
</tr>
<tr>
<td>-same spell out</td>
<td>-different spell out</td>
</tr>
</tbody>
</table>
Monolingual Competition

Studies on polysemy (e.g. Pylkännen et al. 2006) show that different senses of a polyseme have shorter M350 latencies. Berretta, Fiorentino & Poeppel, 2005) show that polysemy is facilitative and homonymy is inhibitive in a LDT.

<table>
<thead>
<tr>
<th>Polysemes</th>
<th>Homophones</th>
</tr>
</thead>
<tbody>
<tr>
<td>e.g. The <em>paper</em> is owned by Murdoch. The <em>paper</em> was written by Elena.</td>
<td>e.g. He fell off the river <em>bank</em>. She opened an account at the <em>bank</em>.</td>
</tr>
<tr>
<td>-faster activation</td>
<td>-slower activation</td>
</tr>
<tr>
<td>-share root</td>
<td>--don’t share root</td>
</tr>
<tr>
<td>-same spell out</td>
<td>-same spell out</td>
</tr>
</tbody>
</table>
Interlingual Allomorphs

What are traditionally called translation equivalents.

Interlingual Allomorphs

e.g. √dog and √chien

- share root (bilingual root competition)
- different spell-out
- polysemy in bilingual speech context
The Role of Phonology

“Translation equivalent primes (both cognate and non-cognates) ---aka interlingual allomorphs-- produce facilitation via their shared meaning representation.”

--Nakayama et al. (2013)

Like polysemy – they share a root
The Role of Phonology

Consider the phonological comparisons translation equivalents between Japanese/English:

– Cognate: /remoN/ ‘lemon’
– Non-cognate: /josei/ ‘woman’

When activating the same root, the phonological overlap facilitates recognition
Phonology & the Lexicon

Phonology is central to understanding the representation and processing of the multilingual lexicon.

DM is the architecture that captures multilingual root competition.
Affixes

Abrahamsson (2003) presents some classic data on the acquisition of coda consonants in L2 Swedish morphology by L1 Mandarin speakers.

His data involve, I would argue, not a functional principle of recoverability, but rather, a DM feature-bundle style analysis.
Swedish ‘r’s

If [r] is part of a lexical stem it will be pronounced more accurately than [r] that is part of present tense or plural affixes.

Retention of an ‘r’ in lexical forms helps recoverability more than the retention of ‘r’ in inflected forms because there are redundant cues to things like tense and plural.
<table>
<thead>
<tr>
<th>Present Tense</th>
<th>Plural</th>
<th>Monomorphemic</th>
</tr>
</thead>
<tbody>
<tr>
<td>kasta-r ‘throw[s]’</td>
<td>sko-r ‘shoes’</td>
<td>dyr ‘expensive’</td>
</tr>
<tr>
<td>gå-r ‘walk[s]’</td>
<td>bil-ar ‘cars’</td>
<td>hår ‘hair’</td>
</tr>
</tbody>
</table>

All the subjects had significantly more errors for multimorphemic words than for monomorphemic words; it’s not just phonology.
The difference in error patterns between inflected versus uninflected forms implicates syntactic features in the explanation.

But what of the differences between performance on [past] versus [plural]?

Remember Abrahamsson invokes a functional explanation: unique markers are retained more than redundant markers.

But, as he admits, it is not easy to tell whether Tense or Plural is more redundant in Swedish.
Prosodic Transfer Hypothesis (Goad & White, 2006; Lardiere, 2007) can’t explain the difference between the suppliance of the two morphemes
Accuracy

The L2 Swedish (L1 Mandarin) subjects are more accurate on Plural than they are on Present.
A DM Transfer Analysis

The singular Number head is null, but plural [+PL] must be marked morphologically. (Yang, 2005).

The underlying plural feature is available for L2 spell out.
‘these students’

```
DP

D⁰
zhe
‘this’,

Num⁰ [+PL]
xie

CL⁰
Ø

nP

xuesheng
‘student’
```
Mandarin doesn’t have Tense but rather Aspect, so (as is well discussed in the literature) there is clearly a learning task here; it does have [finiteness], though Mandarin learners have difficulty with Tense (Hawkins & Liszka, 2003)
Phonology & Morphemes

Phonology is central to the understanding of the spell out of L2 morphemes

DM is the architecture that captures late insertion of L2 morphemes (and intraword codeswitching – Stefanich (in progress))
The Phonetics Interface

🌟 Illusory Vowels
Perceptual Illusions & Phonology

Studies from a number of L1s (Japanese - Dupoux; Matthews & Brown, Korean- Kabak & Idsardi, Brazilian Portuguese – Cardoso; Cabrelli Amaro) reveal perceptual illusions

In production, subjects insert an epenthetic vowel between the obstruents

- baseball → basubaru

Japanese: does not allow obstruent consonantal sequences word medially:

*ac.tor
But this happens in perception too
When exposed to a string like [ebzo], the Japanese subjects *hear* [ebuazo] whether or not there is a vowel present (Dupoux, et al. 1999): Japanese (72% illusory vowel); French (10% illusory vowel)
Thai

- Thai does not allow onset clusters either.
- It does allow medial clusters (like ‘ac.tor’).
- But Thai (unlike Japanese) L1 subjects (since Thai allows medial obstruent strings) do NOT hear an illusory vowel medially (Matthews & Brown).
- When they are presented with [ebzo], they hear [ebzo].
Kabak & Idsardi (2007) show that this phenomenon of vowel epenthesis is mediated by phonological structure (specifically *Coda*) not just by linear adjacency.
sC Onset Perception

There is a cottage industry looking at sC clusters in SLA
## Structural Properties

<table>
<thead>
<tr>
<th>L1</th>
<th>sC Onsets</th>
<th>Branching Onsets</th>
<th>Branching Codas</th>
<th>Errors</th>
</tr>
</thead>
<tbody>
<tr>
<td>Japanese</td>
<td>No</td>
<td>No</td>
<td>No</td>
<td>72%</td>
</tr>
<tr>
<td>Thai</td>
<td>No</td>
<td>No</td>
<td>No</td>
<td>60%</td>
</tr>
<tr>
<td>Brazilian Portuguese</td>
<td>No</td>
<td>Yes</td>
<td>No</td>
<td>50%</td>
</tr>
<tr>
<td>Persian</td>
<td>No</td>
<td>No</td>
<td>Yes</td>
<td>??</td>
</tr>
</tbody>
</table>

This is not a simple task.
### Structural Properties

<table>
<thead>
<tr>
<th>L1</th>
<th>sC Onsets</th>
<th>Branching Onsets</th>
<th>Branching Codas</th>
<th>Errors</th>
</tr>
</thead>
<tbody>
<tr>
<td>Japanese</td>
<td>No</td>
<td>No</td>
<td>No</td>
<td>72%</td>
</tr>
<tr>
<td>Thai</td>
<td>No</td>
<td>No</td>
<td>No</td>
<td>60%</td>
</tr>
<tr>
<td>Brazilian Portuguese</td>
<td>No</td>
<td>Yes</td>
<td>No</td>
<td>50%</td>
</tr>
<tr>
<td>Persian</td>
<td>No</td>
<td>No</td>
<td>Yes</td>
<td>15%</td>
</tr>
</tbody>
</table>

But the Persian subjects are very good!
Why?

- Persian does not allow branching onsets
- Persian allows branching codas up to two consonants
- Persian codas can violate Sonority Sequencing

---

Archibald & Yousefi (2017)
Persian has marked coda clusters (e.g. [tm], [br]) with *rising* sonority (MSD = -5)

English has marked onset clusters (e.g. [st]) with *falling* sonority (MSD = -1)

Codas are more marked than onsets

L2 English is a subset of L1 Persian setting
Redeployment Hypothesis

Based on Archibald (2006)

Persian L1 subjects transfer their marked L1 coda knowledge to be able to acquire the L2 English marked sC onset structures

Both marked strings are characterized by a negative sonority distance
Identification Task
– a forced choice identification experiment

Discrimination Task
– discriminate between /s/ and /es/ word-initally via an ABX discrimination task.
Accuracy of perception and production of /sl/, /sn/, and /st/
Representational Realism

The grammar – not production, not linear adjacency, not direct and accurate storage of the input stream– explains the observed perception patterns

consistent with domain-specific (not general executive cognitive) representations (Blanco-Elorrieta, E., & L. Pylkännen (2016) for language switching)
Production is different

Language control in production recruits *domain-general* regions (dorsolateral prefrontal regions bilaterally) which are also implicated in non-language switching tasks. While perception recruits *language-specific* regions (anterior cingulate cortex) which is not implicated in a non-language switching task.
Illusory Vowels

The perceptual illusions are part of grammar (consistent with Cabrelli Amaro et al., 2017)

The produced epenthetic vowels (i.e. for the Persian L1 speakers), on the other hand, are under cognitive executive control.
Phonology & Perception

- Phonology is central to understanding the phenomenon of illusory vowels in production and perception
- Executive control can be separate from phonology
The Syntax Interface

L2 syntax also reveals the centrality of phonology
Syntax

Languages have two strategies for forming WH questions:

- **English** (Movement): Whom should Bob call?
- **Japanese** (*in situ*): Mito ga nani o katta no?

  Mito-Nom what-Acc bought +Q

'What did Mito buy?'
Richards (2010, 2016) argues that these are two strategies to achieve *contiguity*;

(a) **English**: linear adjacency of C (+Q) and WH

(b) **Japanese**: (i) phonetic boost on the WH element, and (ii) lack of prosodic boundaries between WH and +Q in sentences like (1) compared with (2) where we compare **bolded** objects, and *italicized* minor phrases
Naoya wa なにを にぎやで なんだの？
ナオヤは、何を 飲み屋で 飲んだの？
What did Naoya drink at the bar?

Naoya が なにかを にぎやで なんだ。
ナオヤが、何かを 飲み屋で 飲んだ。
Naoya drank something at the bar.
Operational Question

Will non-native speakers of Japanese show (a) a phonetic boost of the WH words, and (b) a lack of prosodic boundaries between WH and C?
Subjects


The L1 factor did not affect pitch boost data (p=0.7634).

Task

Rehearse in advance, and read out-loud 19 Japanese sentences (WH; Y/N; declarative)

--building on Archibald (2017b)
What did Naoya drink at the bar?
Pitch Boost

What of Richard’s first prediction, that there should be Higher pitch on WH words compared to DPs?

17. Noboru wa piza o mottekitandesu ka?
   Did Noburo bring pizza?

19. Tarō wa nani o mottekitandesu ka?
   What did Taro bring?
## NNS Pitch Average

<table>
<thead>
<tr>
<th>Sentence 17</th>
<th>DP Direct Object</th>
<th>234 Hz Average</th>
</tr>
</thead>
<tbody>
<tr>
<td>Sentence 19</td>
<td>WH Direct Object</td>
<td>228 Hz Average</td>
</tr>
</tbody>
</table>
**NNS Pitch Average**

<p>| | |</p>
<table>
<thead>
<tr>
<th></th>
<th></th>
</tr>
</thead>
<tbody>
<tr>
<td>DP Direct Object All Sentences</td>
<td>208Hz Average</td>
</tr>
<tr>
<td>WH Direct Object All Sentences</td>
<td>201Hz Average</td>
</tr>
</tbody>
</table>
A range of statistical tests (Paired t-tests (p=0.475), Generalized Linear Mixed Effects models, all showed that there was no significant difference between the pitch on WH words and the pitch on DPs.

The NNS L2 phonetic implementation was not nativelike.
Prosodic Structure

What of the second prediction about phonological structure?

On 2 key minimal-pair sentences, subjects, showed no prosodic rises (i.e., no prosodic boundaries) between the WH word and the Question particle.
WH [Word 1  Word 2  Word3] +Q
## Pitch in the Wh-Domain

<table>
<thead>
<tr>
<th>Subject #</th>
<th>Nanio WH</th>
<th>Nomiya-de</th>
<th>Nonda</th>
<th>No [+Q]</th>
</tr>
</thead>
<tbody>
<tr>
<td>1</td>
<td>141 Hz</td>
<td>103 Hz</td>
<td>108 Hz</td>
<td>140 Hz</td>
</tr>
<tr>
<td>15</td>
<td>327</td>
<td>242 Hz</td>
<td>242 Hz</td>
<td>280 Hz</td>
</tr>
</tbody>
</table>
Note the level pitch between WH and C ( [+Q ] ). For these speakers, we posit the following structures (from Richards, 2010):

```
[DP  wh  [DP  [VP]]  C
 (MinP)  (MinP  )
 (MinP  ) ← Wh domain
```

The WH and the C are not separated by prosodic boundaries.

The subjects clearly show a nativelike pattern:

High pitch WH > no phrase boundaries > high pitch +Q
Statistical Results

A Linear Mixed Effects model with speakers as random effects and Word and Sex as fixed effects was fitted to the data.

A second LME showed that Proficiency was not significant; even the Intermediate subjects were good.
However, word position within the Wh-domain (word1, word2, word3) was significant (p = .001 for all comparisons)

There was a significant decline in pitch at each position between the WH and +Q

<table>
<thead>
<tr>
<th>WH</th>
<th>Word 1</th>
<th>Word 2</th>
<th>Word 3</th>
<th>+Q</th>
</tr>
</thead>
</table>
Taken together, these data indicate that the *phonology* of the L2ers is representationally nativelike while the *phonetic* implementation is not.
Consistent with Elfner (2015), these L2 prosodic domains appear to be derived directly from the syntactic structure.

Thus, these data suggest that IL grammars follow the premises of Match Theory (Selkirk, 2011).

L2ers are not transferring L1 structures but are actually acquiring targetlike Syntax.
Note that these IL grammars show evidence of phonological recursion (a hot topic)

The grammars show the phonology mirrors complex syntax – in fact, is a diagnostic of syntactic structure

These are not ‘shallow’ grammars (contra Clahsen & Felser, 2010)
Phonology & Syntax Interface

- L2 learners are able to acquire categorical phonological markers of complex syntax.
- These recursive phonological phrases are not easily read of the input (and are not taught in class).
- Such complex interface properties are acquired by classroom learners (with fairly minimal input) potentially vitiating input-driven accounts.
Summing Up

- No rampant heterogeneity necessary
- Parsimony (Occam) and evidence all lead us to Homogeneity; the Single Engine
- No special machinery for bilinguals
- L2ers learn and process categories and merge them into complex hierarchies
One theme of this conference is bilingual speech.

As L2 speech researchers, we need to recognize the complexity of representation and computation of phonological knowledge that underlies bilingual speech.
L2 phonology is:
– a key component of all modular interfaces;
– a key pillar of GenSLA
these interface phenomena *can* be acquired
And we can meet our epistemological challenges.
A Single Engine; A Common Thread
Thank you.

johnarch@uvic.ca
References


http://linguistics.oxfordre.com/page/recently-published


References


References


Hawkins & Liszka, 2003


Evolution

There is much current discussion on the evolution of human language (Berwick & Chomsky, 2016)

The evolution of the grammatical property of recursion is central (and controversial) – Wolfe (2016); Everett (2005)
Phonological recursion is discussed less often but is just as central.
Prosodic Word Tree
(11) \[ L^*H_P \quad L^*H_P \quad H^* \quad H^*L_P \]
[[[Also jetzt steht]_P [[links]_P [der Gorilla]_P]_P [[rechts]_P [neben dem Gorilla]_P]_P
\quad \text{so now stands \ left \ the gorilla \ right \ beside \ the \ gorilla}
L^*H_P \quad L^*H_P \quad H^*L_P \quad H^*L_I
[\text{das Pferd}_P]_P [[\text{und neben dem Pferd}]_P [\text{rechts}_P [\text{der Löwe}]_P]_P]
\quad \text{the horse \ and \ beside \ the horse \ right \ the \ lion}
\text{‘So now the gorilla is standing to the left; the horse is standing to the right beside the}
\quad \text{gorilla; and the lion is standing beside the horse to the right.’}
[also steht jetzt links der Gorilla]_p1
[rechts neben dem Gorilla das Pferd]_p2
[und neben dem Pferd rechts der Löwe]_p3
The Phonetics Interface II

🌟 Intake Frequency
Input vs. Intake

- **Input:** The linguistic environment
- **Intake:** The subset of the linguistic environment processed by a learner at a given time (Corder, 1967; VanPatten, 1996; Carroll, 2001)
The Challenge

To avoid circularity:

Q: why is it accurate?
A: because it was intake.

Q: how do you know it was intake?
A: because it is accurate.
It’s all out there in the *input* but what becomes *intake* first?
Robust Cues

“a robustly encoded signal is more likely to survive signal degradation or interference in reception”

it is more likely to become intake

what starts as a property of the signal, becomes a property of the representation

by “surviving degradation” a string is more likely to become lexicalized (or phonologized) in the L2
Ejectives

Gonzalez (2014) looks at the acquisition of Yucatec Mayan ejectives by NS of Spanish

Spanish lacks the [constricted glottis] feature

Can they acquire it in L2 Yucatec Maya

AX auditory discrimination task

Forced choice picture selection task
NNS not significantly different from NS in onset position

However they *are* significantly different from the NS in coda position

The recoverability cues for ejectives are much subtler in coda position

Ejectives are robust due (in part) to their dual release bursts (oral & glottal).
Onsets Exploded

Within the onsets, though there are differences in terms of accuracy of perception:

\[ k'/p' > t'/tʃ' > ts' \]
Codas Exploded

And note the (almost reverse) pattern in codas:

\[ tʃ' > ts' > k' > p' > t' \]
not all exemplars of [constricted glottis] are parsed at the same time

Onset, non-strident stop: boost intake frequency

Coda, strident stop: boost intake frequency
These are grounded typologically and phonetically.

Perceptual accuracy paves the way for grammatical restructuring and the phonologicization of [constricted glottis].
Learning is mitigated by intake frequency which is, in turn, modulated by robust transitional cues. Elements which become intake earlier are represented in the lexical entry earlier.
Phonology & Processing

- Phonetic properties are central to understanding the developmental path of acquiring new phonological categories.
- Gradient processing does not necessitate gradient storage.