The redeployment of Persian coda structure in the acquisition of English sC onset clusters: production/perception asymmetries in illusory vowels

John Archibald & Marziyeh Yousefi
Department of Linguistics

Research Question: Why do Persian L1 subjects, whose L1 lacks onset clusters, show highly accurate perception of L2 English onset clusters, yet epenthesize so frequently in production tasks?

It is well-documented that L1Xers evidence epenthetic repair strategies in production tasks of strings that are ungrammatical in the L1.

I speak Spanish

However, it is also clear the this is not the result of a late production routine. Epenthetic vowels also emerge in perception tasks.

Japanese L1ers (Dupoux et al. 1999) when presented with acoustic input of [e] actually hear [e] because of the phonotactics of Japanese.

Illusory Vowels

<table>
<thead>
<tr>
<th>L1</th>
<th>Perception Errors</th>
</tr>
</thead>
<tbody>
<tr>
<td>Japanese (Dupoux et al., 1999)</td>
<td>72%</td>
</tr>
<tr>
<td>Thai (Imsri, 1999)</td>
<td>60%</td>
</tr>
<tr>
<td>Brazilian Portuguese (Cardoso et al., 2007)</td>
<td>50%</td>
</tr>
</tbody>
</table>

Our Subjects
Fifteen native speakers of Persian.

Our Tasks
(1) Identification
10 [st]: 10 [sC] 10 [st]: 10 [se]

"Does the item you will hear begin with a vowel or a consonant?"

Dependent measure: Accuracy

(2) Discrimination
An ABX discrimination task with 800ms ISI.
10 [st]: 10 [sC]: 10 [st]: 10 [se]

"Is the 3rd sound you hear more like the 1st or the 2nd?"

Dependent measure: Accuracy

A comparison of the two tasks showed they did not behave significantly differently (p=.232) so the scores were combined.

High Perceptual Accuracy

The subjects were very accurate on their perception. Why?

Persian Syllables: Onsets & Codas

Persian does not allow branching onsets but does allow branching codas; undeniably, a marked grammar.

E.g. [ma'ar] = 'man'
E.g. [səf] = 'corporation'

Persian Syllables: Marked Codas

Furthermore, Persian allows branching codas with rising sonority.

E.g. [xatm] = 'funeral'
E.g. [oatb] = 'grave'

These are Monosyllables

Different Clusters

Contrary to the Redeployment Hypothesis, [st] clusters were significantly less accurate than [sC] and [st]. P = .001(GLMM) and Odds Ratios over 2.0. There was no difference between the accuracy of [st] and [sC].

• Maybe this is merely a power issue that will disappear with more subjects.

• Maybe the subjects are actually treating these string as codas, and (following Kaye (1992), Goad (2016), and Enochson (2014) assigning the [st] to the coda of an already-headed syllable. [st] is the worst syllable contact (Vennemann, 1987).

Underlying Representations

The illusory vowel data present challenges for models which assume that the underlying representation is always a mirror of the input and that the output is the focus of critical data.

Production

The production data show frequent epenthesis even in light of accurate perception. This is consistent with recent work on language switching (Blanco-Elorrieta, E., & L. Pylkännen, 2016) which shows that production is mediated more by domain-general executive control than domain-specific grammatical representation. The Advanced subjects were at 90% accuracy.

Conclusion

Persian L1 subjects are more accurate in the perception of English L2 onset clusters than other L1Xers who lack sC onsets because L1 Persian can transfer the L1 coda MSD which licenses the L2 onsets. The production/perception asymmetries argue for representational realism.

References


Archibald (2005). Lardiere (2009) demonstrate that L1Xers can use L1 building blocks to assemble new LX structures. The Persian L1ers can redeploy their L1 coda MSD knowledge to the L2 onsets as all English onset sequences will be licensed. Redeployment would predict high accuracy but no differences between strings.

The Redeployment Hypothesis

Archibald (2005). Lardiere (2009) demonstrate that L1Xers can use L1 building blocks to assemble new LX structures. The Persian L1ers can redeploy their L1 coda MSD knowledge to the L2 onsets as all English onset sequences will be licensed. Redeployment would predict high accuracy but no differences between strings.
