The Redeployment of Marked L1 Persian Codas in the Acquisition of Marked L2 English Onsets: Redeployment as a Transition Theory

John Archibald & Marziyeh Yousefi,

Dept. of Linguistics, University of Victoria,

ConCALL 3
Indiana University

University of Victoria

Illusory Vowels

- 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
  - strike ➔ suturiku
- 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* [ebɯzo] whether or not there is a vowel present (Dupoux, et al. 1999): **Japanese** (72% illusory vowel)versus **French** (10% illusory vowel)

- **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 (Carlisle; Yavas & Sommeilan, 2010)
- Brazilian Portuguese (Cardoso):
  - Does not allow sC clusters
  - Allows Obsruent + Liquid clusters (e.g. [br])
  - Allows maximally single C codas
- Both production and perception studies showed definite differential accuracy effects (and no ceiling effects):
  - Production: sl > sn > st
  - Perception: st > sl > sn

- The Brazilian Portuguese L1 subjects had difficulty (performing at chance) discriminating accurately between forms which began with:
  - sC and isC

(where [i] is the BP epenthetic vowel)

- The same is true in Thai (Imsri):
  - No sC onsets
  - No branching codas
- In *production*, they epenthesize to break up the sC:
  - spa ➔ səpa
- In *perception* the advanced learners made 60% errors on discriminating sC strings from SVC strings
  - Even when correct, there were significantly longer RTs
  - And remember, they did fine on [ebzo]
- So this is mediated by *grammar*

Our Languages

- L1 = Persian
- L2 = English

- Recent data from Yousefi (2017) suggest that Persian speakers (who also lack sC onset clusters) do *not* exhibit such perceptual illusions
- Even though they have been documented to epenthesize in production (Karimi, 1987; Yarmohammadi, 1995)

The Tasks

- Perception
- Production

Perception

- Identification Task
  - a forced choice identification experiment
- Discrimination Task
  - discriminate between /s/ and /es/ word-initially via an ABX discrimination task.

Identification

- 10 [st]; 10 [est]
- 10 [sn]; 10 [esn]
- 10 [sl]; 10 [esl]

- “Does the item you will hear begin with a vowel or a consonant?”

Discrimination

- An ABX discrimination task with 800ms ISI.
- 10 [st]; 10 [sn]; 10 [sl]

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

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

Production Tasks

- **Formal Task**
- 29 sentences in all, the target clusters /st/, /sn/, /sl/ occurred 10 times for each cluster
- *Example*: *Instructions*: Read aloud the following sentences, please.

  Dan slept early today

Production Tasks

- **Informal Task**
- 12 pictures consisted of 3 words for each cluster (i.e. 3 /sn/, 3 /st/, and 3/sl/) as well as three distracters
- *Example*: Pictures of the item “slippers” and ‘stars” in the informal production task.

- The tasks were not significantly different (p=.133)

The Subjects

- Round 1: 15 NS of Persian
- Round 2: additional 5 NS of Persian (perception only)

Data Patterns (Perception)

L1	sC Onsets	Branching Onsets	Branching Codas	Errors
Japanese	No	No	No	72%
Thai	No	No	No	60%
Brazilian Portuguese	No	Yes	No	50%
Persian	No	No	Yes	??

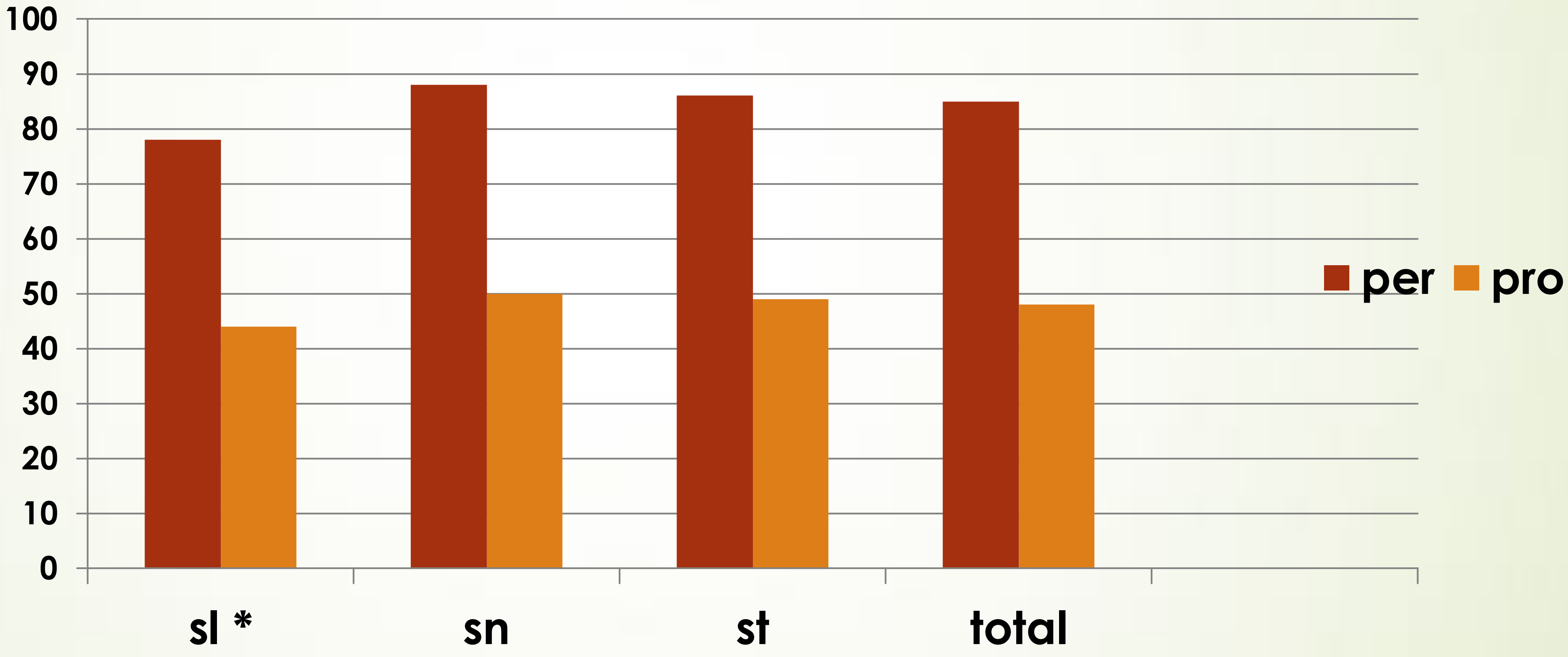
L1	sC Onsets	Branching Onsets	Branching Codas	Errors
Japanese	No	No	No	72%
Thai	No	No	No	60%
Brazilian Portuguese	No	Yes	No	50%
Persian	No	No	Yes	**14%**

- Even the Beginner students scored *75% accuracy* (compared with Cardoso's BP Beginner's who performed at *chance*, and Matthews & Brown's (2004) Thai subjects who made *60% errors*).
- Thus, we note that the Persian subjects are very accurate in perceiving the L2 sC onsets

Production

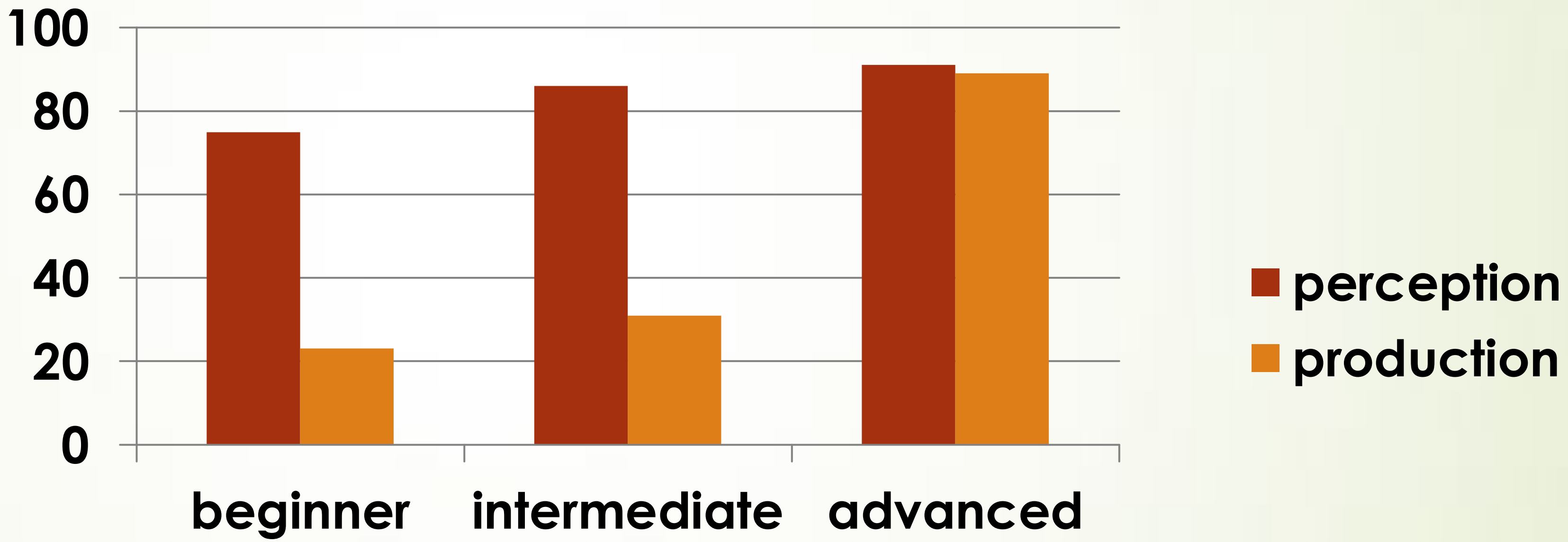
- Even though they perceive accurately, they still produce epenthetic vowels

Accuracy X Cluster Type



Note the very high accuracy rates in perception.

Accuracy X Proficiency



- Perception was significantly more accurate than production ($p=.004$), though they were correlated (Pearson $r = .536$).

Redeployment I

- Archibald (2005) for phonology, and Lardiere (2009) for morphosyntax demonstrate that L2ers can use L1 building blocks to assemble new L2 structures. The Persian L1ers can redeploy their L1 coda MSD knowledge to the L2 onsets where *all* English onset sequences will be licensed.
- Redeployment would predict high accuracy and no differences between strings.

- Persian allows coda clusters which violate the SSP (in monosyllabic, monomorphemic forms) with *rising* sonority. Some examples are:

  - *xætm* ‘funeral’ (Sonority Distance = -2)
  - *qæbr* ‘grave’ (Sonority Distance = -3)

Persian Codas

Persian Sonority Distance in Codas (data from Kambuziya & Zolfaghari 2006, - means fall and + means rise in sonority)

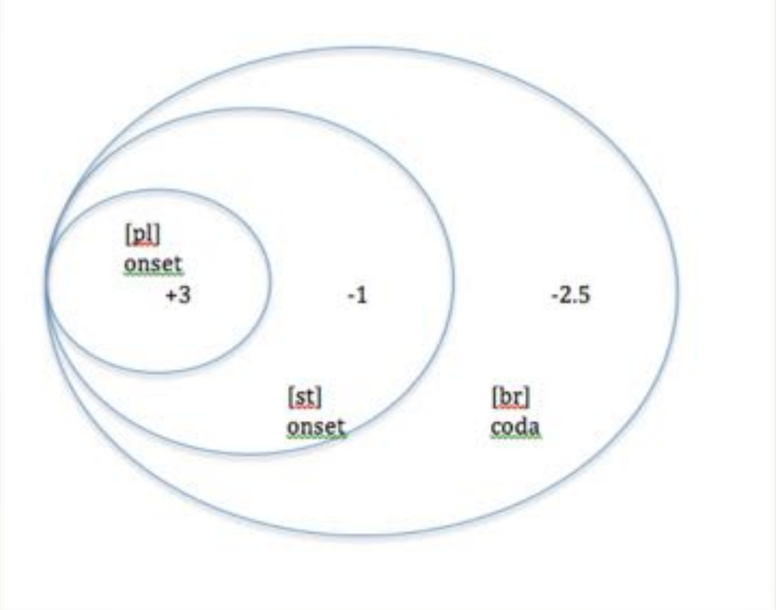
SD	-4	-3	-2	-1	0	+1	+2	+3
Strings	yb yd	rk rd lb	nd rs ng	St ʃt ft xt nz sk ʃk	tb bʔ mn fʃ tk sf sx Sh fv fz	bh bx bʃ ɢʃ dv ds bz mr zm Sn zn ml xm	zl fl sl xl tn fr sr ʃr km hr	br dr kl tr
words	eyb eyd	xark Kard qalb	kand Fars rang	rast daʃt raft rixt tanz usk	ɢotb tæbʔ ʔæmn kæfʃ potk	æbx næbʃ hæml zebh sobh	fæzl tefl ɢosl boxl mætn kofr mesr	bæbr zebr sedr ɢædr ʃekl

Redeployment predictions on the acquisition of the sC clusters:

•Persian learners of English can redeploy their L1 coda MSD knowledge to the L2 onsets where all English onset sequences will be licensed.

•Redeployment will predict high accuracy but no differences between strings

Knowledge of Persian MSD and branching codas redeployed in learning English



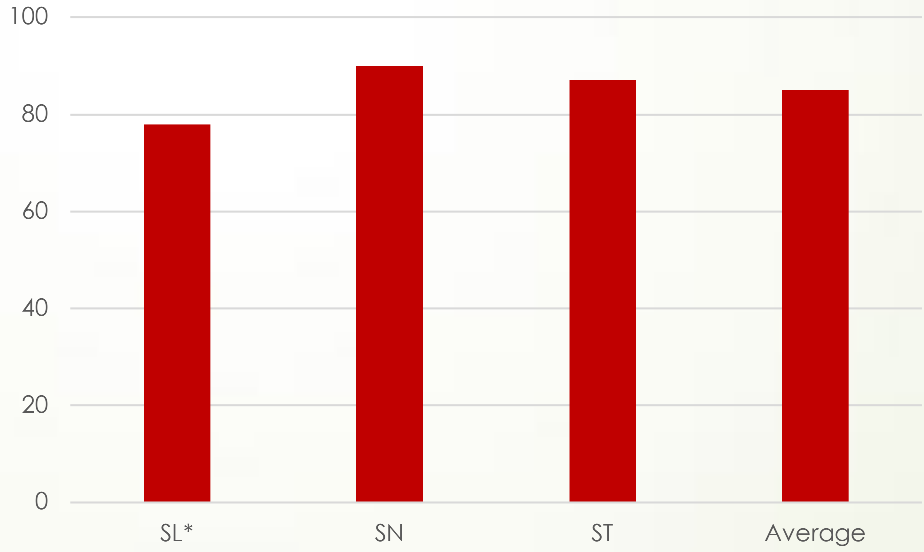
Hypothesis: Predicted path based on property redeployment theory: /sl/=/sn/=/st/

Results

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

Additional Subjects

- We ran 5 more subjects bringing n to 20 to see if the [sl] effect would disappear
- It didn't



GLMM table

```
## Fixed effects:
##                Estimate Std. Error z value Pr(>|z|)
## (Intercept)      0.9097     0.6845   1.329 0.183869
## ConsonantSN      0.8201     0.2355   3.482 0.000497 ***
## ConsonantST      0.6066     0.2277   2.664 0.007731 **
## Level2           0.9938     0.8997   1.105 0.269359
## Level3           2.6680     0.9574   2.787 0.005323 **
## ---
## Signif. codes:   '***' 0.001 '**' 0.01 '*' 0.05
```

Extant Cluster Predictions

- Markedness: sl > sn >st
- Frequency: st> sl > sn

Nobody predicts that [sl] will be the worst!!

Except.....Syllable Contact (Preference Laws)

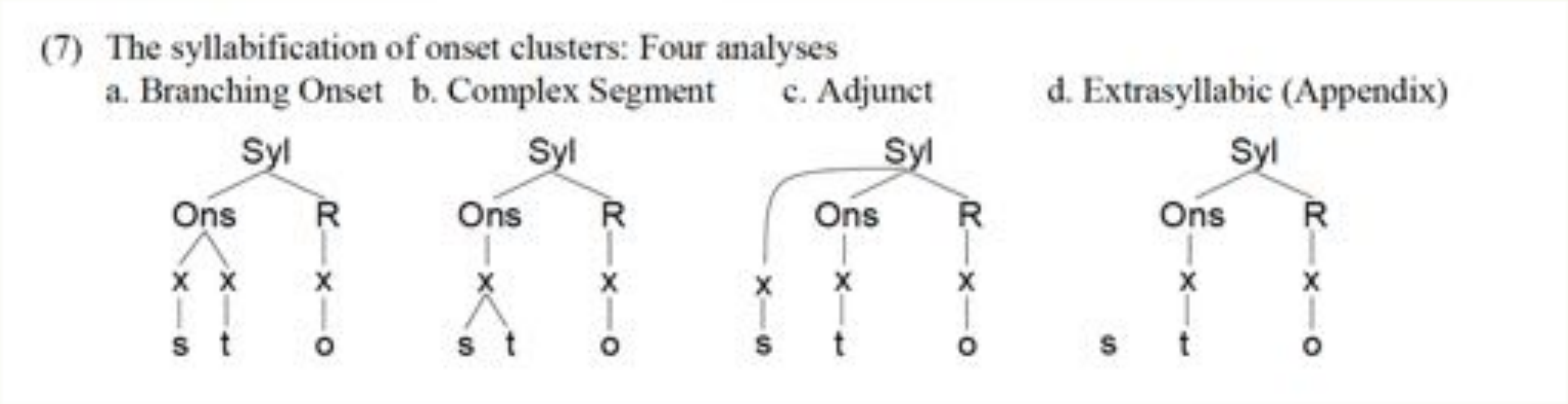
- Murray & Venneman (1983); Enochson (2014)
- The greater the sonority drop from coda to onset the more harmonious the contact. So [s . t] is a good contact and unlikely to be repaired while [s . l] is the worst contact (of our three) and most likely to be repaired.

Syllable Contact

- Maybe the subjects are actually treating these strings as *codas*, and (following Kaye (1992), Goad (2016), and Enochson (2014) assigning the [s] to the coda of an empty-headed syllable. [s.l] is the worst syllable contact (Vennemann, 1987).
- Enochson (2014) showed production accuracy of:
  - [st] > [sn] > [sl]
  - 86% 79% 60%

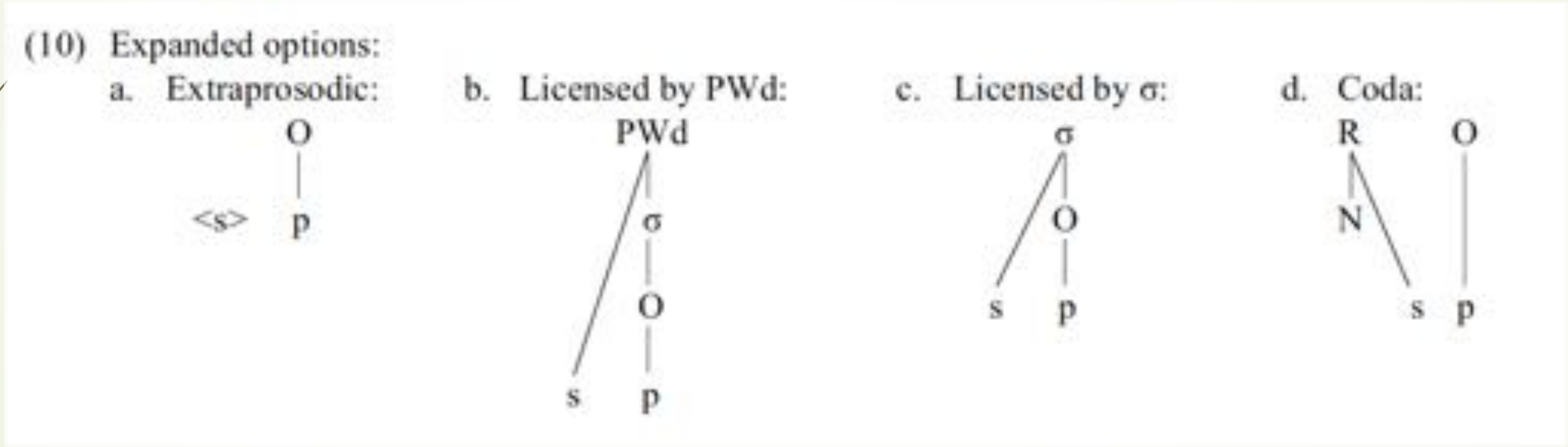
English Left-Edge

From Cardoso (2007)



English Left-Edge

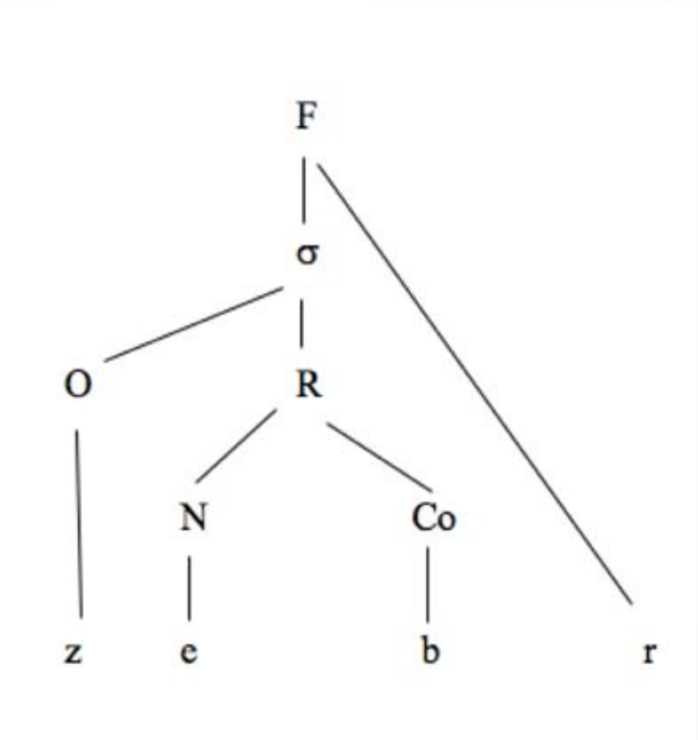
From Goad (2016)



Persian Right-Edge Appendices

- Persian syllables have maximally two C's at the right edge; thus, an appendix

The image part with relationship ID rId2 was not found in the file.



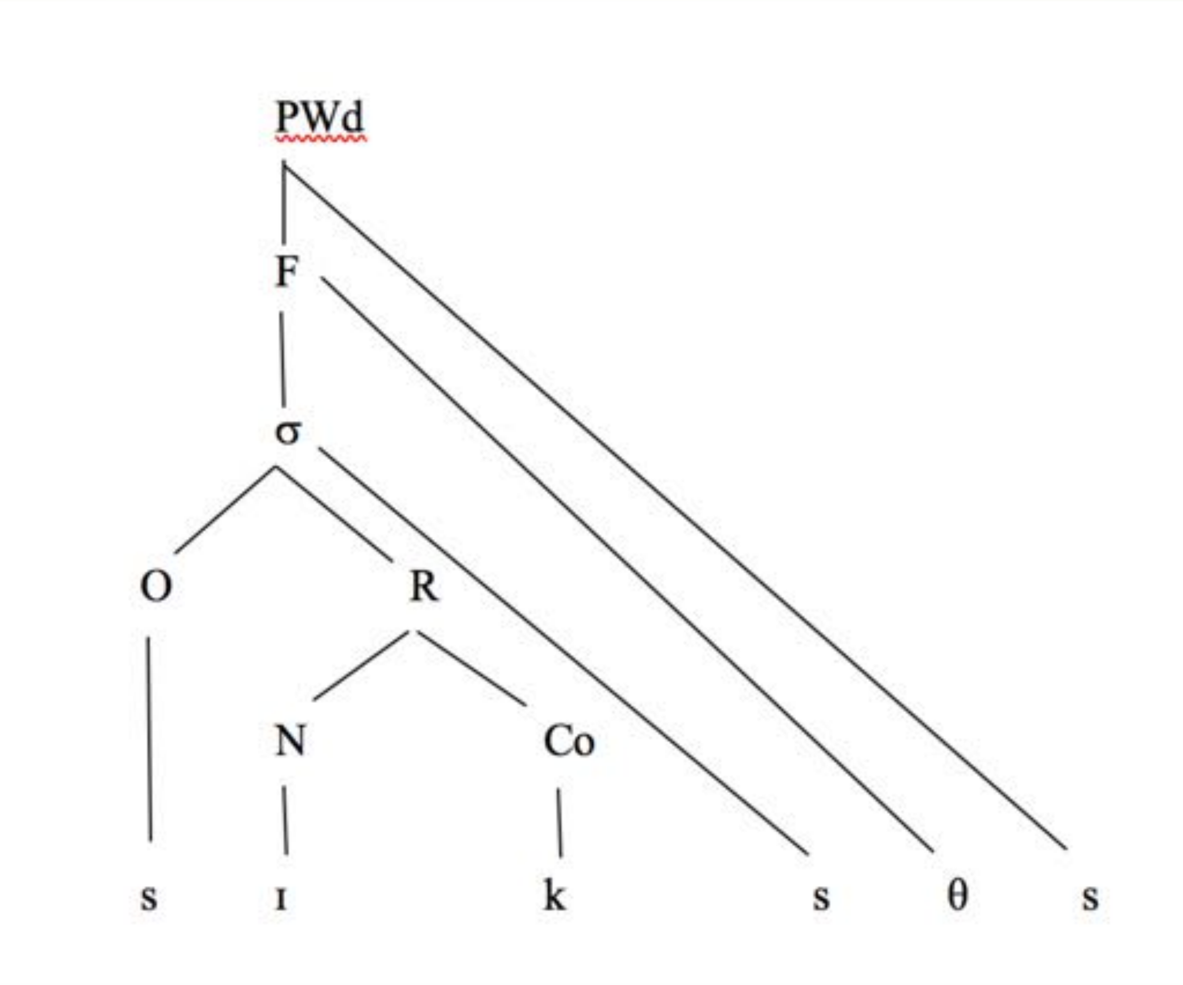
Connecting Onsets & Codas

- Davis & Baertsch (2010) adopt the Split Margin model to capture a structural relationship between onsets and codas
- Yousefi (2017) proposed that the Persian coda MSD of -3 would transfer to English and license all English CC sequences (even [st])

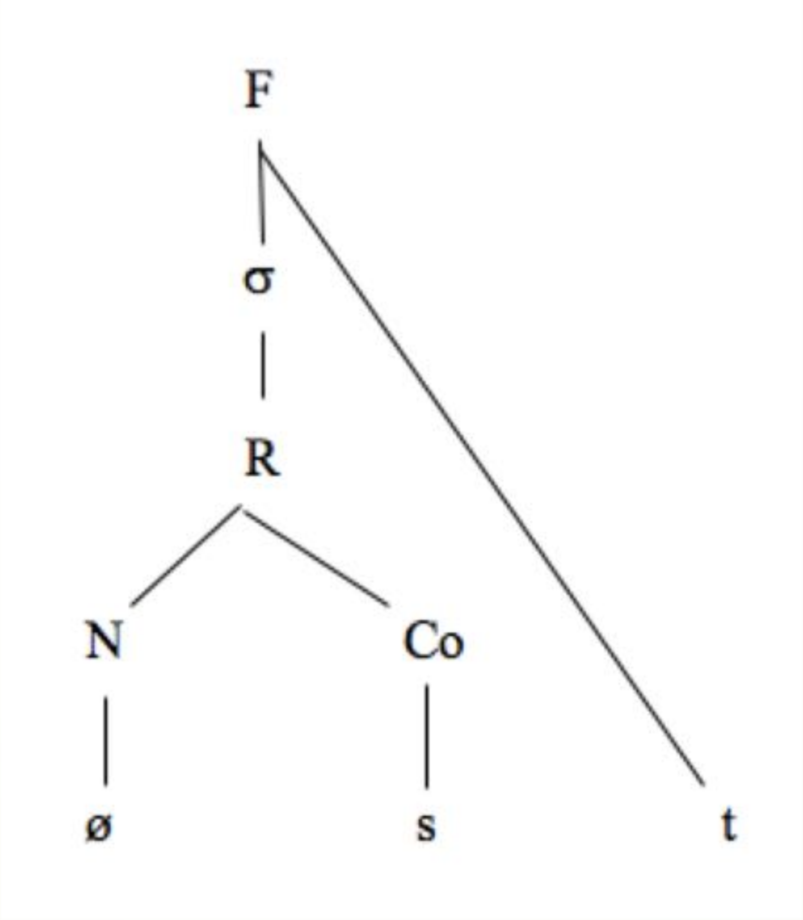
'Branching Codas' Revisited

- Up til now we've been casually referring to 'branching codas' to describe our data
- But most phonological theories do not sanction branching codas (Golston & Kehrein (2004); Kiparsky (2002); Vaux & Wolfe (2009); Watson, J. (2007))
- So, let us explore our theoretical account

English Right-Edge Appendices



Initial Transfer from Persian



- Given Persian right-edge structures, we assume that the entire [sC] sequence is initially assigned to the first syllable. As proficiency increases, the learners will resyllabify the 2nd C to the following onset. The primary reason for the slower acquisition of the [sl] onset has to do with L2 input frequency. It takes the subjects longer to acquire the [sl] cluster because it is less frequent in the L2 input (Cardoso, 2007) – the difference between [sl] and [sn] is very small

- Then [t] becomes optimal onset and [s.t] becomes optimal syllable contact
- Then [n] becomes an allowable onset
- Then [l] becomes an allowable onset

- It takes TIME to overcome the sub-optimal contact; they need positive evidence to make those changes. The positive environmental evidence is less available.

- This structure explains why the sC clusters trigger prothesis while the [pl] (and all other) clusters trigger epenthesis (Fatemi et al., 2012; Fleischhacker, 2001; Karimi, 1987).
- E.g., p[e]lastic versus [e]smoke

Redeployment Redux

- As a Property Theory
- As a Transition Theory
    - From R. Cummins (1983).

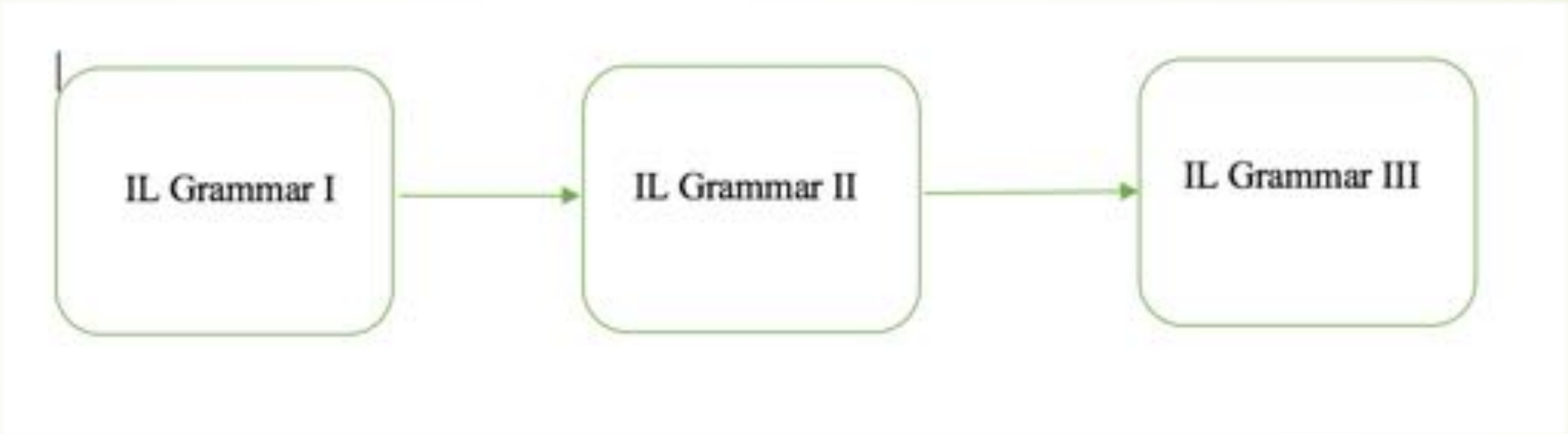
Property Theories

- Theories of synchronic knowledge
- What is the initial state?
- What is the end state?


L1 Grammar
IL Grammar
L2 Grammar

Transition Theories

- Theories of representational change
- Developmental path


IL Grammar I
IL Grammar II
IL Grammar III

Why the Difference?

- We propose that the performance of all the subjects is explained, in part, via properties of their L1 appendix structure
- Japanese, Thai, BP transfer their L1 knowledge and do *not* have the building blocks to handle sC onsets and the perceptual illusion of vowel insertion occurs;
- The illusory vowel is actually part of their stored representation

- But the Persians seem to be able to set up accurate underlying representations because of the L1 grammar

- Persian subjects *redeploy* their L1 knowledge of post-vocalic CC strings to their perception of L2 sC strings thus overcoming the perceptual illusion
- They learn *quickly* that the illusory vowels are not part of the stored representation
- They have appendices in their L1 and transfer this to the L2

English vs Persian

- Persian fills empty-headed syllables
- English does not

- Parametric variation?
- Constraint ranking?
  - Nodes must be filled (DEP)
  - English can violate DEP but Persian can't; so DEP is higher ranked in Persian and needs to be demoted in L2 English
- We remain agnostic here

- The Persian learners of English learn that English doesn't need to fill empty nuclei
- They hear this easily in perception (intelligible) though in production they are still producing them
- Their production is not a direct mirror of their underlying/input representation.
- Persian inputs are nativelike; Japanese inputs aren't.

Architectural Implications

- These illusory vowel data present challenges for models which assume (a) that the underlying representation is always a mirror of the input and (b) that the output is the locus of critical data.
- How could Harmonic Serialism which views changes to be driven serially to improve the harmony of the output form handle these kinds of data?

Summary & Conclusions

- Even L1s with CC Onsets can have difficulty perceiving English sC sequences
- Yet L1s with right-edge appendices (and no CC initial strings) are able to accurately perceive L2 English sC initial sequences
- The accuracy is explained is the L2 target is a coda+appendix string with a null nucleus as the potential site for a prothetic vowel in production
- Redeployment (plus syllable contact laws) explains the developmental path and the difficulty with [s.l] strings
- The construct of intelligibility can be rethought within parsing theory and models of spoken word recognition

Thank you

johnarch@uvic.ca

References

- Archibald, J. (2006). Second language phonology as redeployment of phonological knowledge. *Canadian Journal of Linguistics* 50 (1/2/3/4). Pp 285-314.
- Archibald, J. (2004). Interfaces in the prosodic hierarchy: New structures and the phonological parser. *International Journal of Bilingualism* 8(1): 29-50.
- Archibald, J. (2003). Learning to parse second language consonant clusters. *Canadian Journal of Linguistics* 48(3/4): 149-178.
- Atkey, S. (2001). *The Acquisition of Czech Palatal Stops*. MA Thesis. University of Calgary.
- Boudaoud, M. & Cardoso, W. (2009). The variable acquisition of /s/ + consonant onset clusters in Farsi-English interlanguage. In M. Bowles, T. Ionin, S. Montrul, & A. Tremblay (Eds.), *Tenth Generative Approaches to Second Language Acquisition Conference* (pp. 86-104). Somerville, MA: Cascadilla Proceedings Project.
- Cabrelli Amaro, J. et al. (2017). Phonotactic restructuring in L1 Brazzilian Portuguese. Paper at ISB.
- Cardoso, W., P. John, & L. French (2007).The variable perception of /s/+Coronal onset clusters in Brazilian Portuguese English. In *Proceedings of New Sounds 2007*. Pp. 86-106

- Cardoso, W. (2007). The development of sC onset clusters in interlanguage: markedness versus frequency effects. In *Proceedings of GASLA*. Cascadilla Press.
- Carlisle, R. (2006). The sonority cycle and the acquisition of complex onsets. In B.Baptista & M. Watkins (Eds.), *English with a Latin beat: Studies in Portuguese/Spanish English interphonology* (pp. 105-137). Amsterdam: Johns Benjamins
- Dupoux, E., K. Kakehi, Y. Hirose, C. Pallier, & J. Mehler (1999). Epenthetic vowels in Japanese: a perceptual illusion? *Journal of Experimental Psychology: Human Perception and Performance* 25: 1568-1578
- Enochson, K. (2014).L2 production of English onset sC and CC clusters. Concordia Working Papers in Applied Linguistics; Proceedings of New Sounds 2013. Pp. 171-184.
- Fatemi, M.A., A. Sobhani, & H. Abolhassani (2012). Difficulties of Persian leaners of English in pronouncing some English consonant clusters. *World Journal of English Language* 2(4): 69-75
- Fleischhacker, H. (2001). Cluster-dependent epenthesis asymmetries. *UCLA Working Papers in Linguistics, 7*, 71-116.

- Goad, H. (2016). Phonotactic evidence from typology and acquisition for a coda+onset analysis of initial sC clusters. Kim et al., eds. *Proceedings of 33rd WCCFL*. Cascadilla Press. Pps. 17-28.
- Imsri, P. (1999). Thai speakers' perception and production of English onset clusters /sC-/. Unpublished manuscript. University of Delaware.
- Iwasaki, S. & P. Ingkaphiram (2005). *A Reference Grammar of Thai*. Cambridge
- Jabbari, A. A. & Samarvarchi, L. (2011). Persian learners' syllabification of English consonant clusters. *International Journal of English Linguistics, 1*(1), 236-246.
- Karimi, S. (1987). Farsi speakers and the initial consonant clusters in English. In G. Ioup & S. H. Weinberger (Eds.), *Interlanguage phonology: The acquisition of a second language sound system* (pp. 305-318). Cambridge, MA: Newbury House.
- Kaye, J. (1992). Do you believe in magic? The story of s+C sequences. *SOAS Working Papers in Linguistics & Phonetics: 293-313*.
- Kehrein, W. and C. Golston (2004). A prosodic theory of laryngeal contrasts. Phonology 21: 325-357.
- Matthews, J. & C. Brown. (2004). When language intake exceeds input: language specific perceptual illusions induced by L1 prosodic constraints. *International Journal of Bilingualism* 8(1): 5-27.
- Munro, M. & T. Derwing (1995). Foreign accent, comprehensibility and intelligibility in the speech of second language learners. *Language Learning* 45: 73-97.

- Munro, M., T. Derwing, & S.L. Morton (2006). The mutual intelligibility of foreign accents. *Studies in Second Language Acquisition* 28: 111-131.
- Summerell, F. (2007). *The L2 Acquisition of Japanese Length Contrasts*. MA Thesis University of Calgary.
- Yarmohammadi, L. (2005). *A contrastive phonological analysis of English and Persian: A course book in applied phonological studies.* Shiraz, Iran: Shiraz University Press.
- Yavas, M. & M. Someillan. (2010). Patterns of acquisition of /s—clusters in Spanish-English bilinguals. *Journal of Multilingual Communication Disorders* 3(1): 50-55.
- Yavas, M. (2006). Sonority and the acquisition of #sC clusters. *Journal of Multilingual Communication Disorders, 4(3),* 159-16.
- Yousefi, M. (2016). Perception and production of sC onset clusters in Persian speakers of English. Draft PhD candidacy paper. University of Victoria.

Future Predictions

- Gonzalez (2004) argues that Spanish has a Foot-level appendix for [s] at the right edge.
- This predicts (for us) that Spanish PERCEPTON of sC onsets should be good - is it?
- It predicts that Arabic dialects with Right-Edge appendices (e.g. Hijazi) should do as well as those with branching onsets (e.g., Najdi); they do (Alhemaid, in progress). This confirms (consistent with BP data) that the locus of explanation is the appendix not the branching onset

‘Hearing’ sC

- Who ‘hears’ sC sequences?
- L1s with appendices find English sC intelligible (Munro & Derwing, 1995); Persian 85% accurate
- L1s without appendices find them unintelligible; Japanese 28% accurate; BP 50% accurate

Intelligibility as *Parsability*

- Munro & Derwing refer to intelligibility of L2 speech as the property of allowing the listener to recover the target lexical item
- However, intelligibility is not a property only of the signal
- It is the result of the listener *parsing* the input
- M & D recognize the need to explore listener factors more (and, indeed, do so in Munro, Derwing & Morton (2006)

Intelligibility as *Parsability*

- But we can place intelligibility within the context of lexical processing and spoken word recognition; hence *parsing*
- So, Persian ears parsing sC strings will do different things than Japanese or BP ears parsing sC strings
- The sC strings are intelligible (i.e., parsable) to Persian L1 subjects
- (see Archibald, 2003, 2004 for more details on phonological parsing)

- Intelligibility will be affected more by left-edge effects
- More false hits, more lexical competitors will be activated
- When the Persian subjects hear an English word like *sleep*, they can parse it because of the transfer of L1 structures
- When BP subjects hear a word like *sleep*, they cannot parse it initially, but must learn to parse it

- The crux is to account for an [sl] string assigning the [s] to the coda (of a null-nucleus) while a [pl] string assigns the [l] to an onset and leaves the [p] unsyllabified
- We will not go into the details here but it requires a dependency relation which dictates that [s] goes into the coda and that the following C is licensed as an appendix iff the coda is [s]