

## Corpus-External Evidence

Phonological competence can be revealed in various linguistic behaviors.

- Foreign accent in second language acquisition (interlanguage)
- Loanword adaptation
- Language games (ludlings) and speech disguises
- Orthographic practice
- Judgments of rime, similarity, novel forms
- Language deficits: slips of tongue and aphasia

[1]. Speakers of the same native language acquiring a second language (e.g. English) typically make the same errors. They can be attributed to interference from native grammar and provide evidence for it. However, recent research indicates that some errors cannot be fully explained in terms of the native grammar and require reference to some other component of linguistic knowledge (UG).

- Segment: [y] as ju: French fumer
- Sequence: Spanish: estop, eSteve,
- Rule: Russian drugs [draks]; Arabic sixty -> sikisti
- Mystery: Mandarin learners of English codas. Broselow, Chen, and Wang 1998.

	<u>final voiceless stops</u>	<u>final voiced stops</u>
correct	.19	.02
epenthesis	.36	.36
deletion	.46	.43
devoicing	---	.19

	<u>monosyllables</u>	<u>disyllables SW</u>	<u>disyllables WS</u>
correct	6	13	3
epenthesis	43	21	18
deletion	5	75	32
devoicing	6	11	7

- no final voicing
- preference for SW stress sequence
- Emergence of Unmarked: in the transitional inter-language unmarked structures emerge

[2] loans in Japanese

- segment : coffee > ko:hi:
- sequence: CV, V, CVX, VX X = first have of geminate, homorganic nasal, N  
spoon > supun, ski > suki
- rule: devoicing s[ʌ]kiyaki; s[ʌ]ki

- mystery: strike > sutoraiku; test > tesuto; in native Japanese /tu/ realized as affricate [t<sup>s</sup>u]. When word is borrowed with /tu/ there is conflict: remain faithful to source vs. native system: earlier loans have affricate while later ones have [tu]. But when vowel is epenthetic, then [tu] much more uniformly avoided in favor of [to]. Since vowel has no source to be “faithful” to, the native restriction against [tu] can be respected at the cost of a more “salient” epenthetic vowel. The epenthesis rule has no precedent in the grammar of Japanese. It is adjustment the Japanese speaker makes to conform to the syllable structure constraints in his grammar.

/tuna/	<u>Faith-u</u>	<u>Faith-t</u>	<u>*tu</u>		
tuna			*		
tona	*				
tsuna		*			
/test/	<u>Faith-u</u>	<u>Faith-t</u>	<u>*tu</u>	<u>*o</u>	<u>*u</u>
tesutu			*		
tesuto				*	*
tesoto				**	

### [3] language games

Chao 1934: Indeterminacy in the phonemicization of Mandarin: the palatal affricate tʃ is found only before the high front vowels [i,y] while k, t<sup>s</sup>, and t<sup>s</sup> (retroflex) are found before all vowels except [i,y]. The affricate [tʃ] is thus in complementary distribution with each of k, t<sup>s</sup>, and t<sup>s</sup> and the question is what phoneme to group it with. Phonetic similarity might exclude the retroflex but what about the alveolar vs. velar? Chao cites two facts in favor of grouping with the velar: native speaker judgment of alliteration: [kə, tʃi, ku, tʃy] are an “alliterative series with only different vowels”. A language game reparsing Onset+Rhyme to Onset –aik-Rhyme shows the application of the palatalization rule: cf. pei -> pai-kei and mi -> mei-tʃi. But this does not tell us what would happen if [i] were placed after an alveolar or a retroflex; a more interesting datum would be the game applied to [tʃi]. Chao’s analysis predicts k-aitʃi.

Cuna (Panama) has various language games documented by the anthropological linguist Joel Shertzer. One is a reversal:

<u>Cuna</u>	<u>Backwards</u>	
i.na	na.i	medicine
da.ge	ge.da	come
sa.ban	ban.sa	belly
ob.sa	sa.ob	bathed
ar.gan	ga.nar	hand
in.na	na.in	chicha
go.e	e.go	deer

In Cuna the voiceless stops [p,t,k] derive from underlying geminates transcribed as /bb, dd, gg/: this analysis is supported by various facts of distribution (e.g. no initial or final clusters, no initial or final p,t,k) and alternation (neg ‘house + gine ‘inside’ -> nekine ‘inside the house’).

<u>Cuna</u>	<u>Backwards</u>	
sa.pan	ban.sab	firewood
sa.te	de.sad	no
da.ke	ge.dag	see

But some speakers have ban.sa, de.sa, ge.da. This would follow from applying the reversal after degemination plus devoicing in order to respect the constraint against initial voiceless stops. Such an interpretation depends on the independence of phonotactic constraints and rules.

[4]Orthographies typically abstract away from recording predictable, non-distinctive features, arguably because native speakers cannot hear them easily. This was strong motivation for the phonemic level in American Structuralism. Imagine trying to spell the flap. Illustrations found in work of Sapir’s students such as Alex Thomas (Nootka): consonants geminated in V\_\_V: but this is never reflected in spelling in contrast to geminates that arise from morpheme concatenation.

#### Japanese Hiragana

- rendaku voicing: kami ‘paper’, ori-[g]ami; yu ‘hot water, toofu ‘tofu’, yu-[d]oofu ‘boiled tofu’
- palatalization and affrication: /ti/ -> [tʃi], /si/ -> [ʃi]; /tu/ -> [tʃu]; for voiced obstruents no fricative—affricate contrast: /di/ -> [dʒi]; /zi/ -> [dʒi]; /du/ -> [dʒu]; /zu/ -> [dʒu].
- in Hiragana syllabary each CV combination registered with a different symbol.
- voiced obstruents [d], [z], [g] are represented as diacritic variants (dakuon) of [t], [s], [k]
- the characters for voiced variants of /ti/ and /si/ are both pronounced [dʒi] and the voiced variants of /du/ and /zu/ are both pronounced as [dʒu].
- the question arises of when to represent [dʒi] as voiced /ti/ or voiced /si/ and when to represent [dʒu] as voiced /tu/ or as voiced /su/. The policy is that the fricatives voiced /si/ and voiced /su/ are used in general except in the following case where the stops voiced /ti/ and voiced /tu/ are used.
- in a compound where the second element is /ti/:

hana ‘nose’, ti ‘blood’: hana- dʒi ‘nosebleed’ spelled with voiced /ti/  
kan ‘can’, tume ‘cram’: kan-dʒume ‘canned food’ with voiced /tu/

- reduplicated: ti- dʒi -mi ‘to shrink’; tu-dʒu-ku ‘to continue’



	<u>complex onset</u>			
unmarked	OLV	tra	2.2%	26/1161
	OGV	tja	6.7	58/865
	NGV	mja	14.4	15/104
	LGV	rja	34.4	136/398
	ONV	tna		
	NLV	mla		
	OLGV	trj	33.0	7/21

Observations:

- complex onsets simplified more than simplex ones
- markedness trends supported in complex onset reduction