24.961   Syllable-2   Syllable parsing

[1] Evidence for syllable divisions
   •   Phonotactic constraints
   •   Prosodic distinctions
   •   Speaker intuitions

Example 1: English tl vs. tr
   •   trip, trap, true vs. *#tl
   •   [ə].tʰrócious vs. [æ]tˀ.lántic

Example 2: Cairene Arabic vs. Spanish
   •   In Cairene Arabic words do not begin with consonant clusters
   •   Clusters attract stress just like a long vowel: kátabu 'they wrote', kitáabu 'his book', katáb-na 'we wrote'
   •   Triconsonantal clusters broken by epenthesis: wálad 'boy', walád-ha 'her boy'; bint 'daughter', bint-áha 'her daughter'
   •   In Spanish post consonantal stop-liq is possible: sem.blanza, claus.tro
   •   stop-liq possible initially: tres, blanco
   •   antepenultimate stress is possible with open-syllable penult: jíbaro 'peasant', fú.ne.bre, idó.la.tra

[2] Donca Steriade (1999, 2012) proposes that restrictions on sound sequencing and location (phonotactics) are string based and not based on syllable divisions
   •   Syllable divisions are not observed but inferred[1]; one source of evidence is permissible word beginnings and endings
   •   "syllables bear structural similarity to the edges of better known constituents such as words"
   •   articulatory correlates to onset vs. coda consonants may exist (e.g. Maddieson 1984 finds that vowels in open syllables are phonetically longer than in closed syllables in most languages) but it has not been shown that speakers perceive these properties and rely on them to learn or parse the string
   •   There is no articulatory correlate to speakers who parse *lemon* differently (lem.on vs. le.mon) and no isoglosses for syllable division

[3] Word-Based syllable hypothesis

---

[1] No abstract linguistic categories are literally observed; what is intended is how direct and reliable the inductive inference is

- Speakers *infer* syllable boundaries by assuming that syllables are one-vowelled words and basic similarity is the segmental composition of the word edges to the syllable edges
- When multiple guidelines conflict then speakers are uncertain and produce variable judgments; where speakers are certain then the guidelines converge on a single answer
- There is a UG preference for V.CV (Onset, No-coda) but in *lemon* the parse [le.mon] conflicts with the fact that no words end in a lax nonschwa vowel such as ε
- so there is a conflict between Onset and word edge criteria and the speaker cannot consult outputs of other speakers since the difference is not perceptible
- hence more uncertainty about the parse of *lemon* vs. *demon* where de.mon is consistent with both V.CV Onset preference and syllable = possible word: [dij] + [mon]

[4] in some cases where evidence from edges is conflicting, it leads to uncertainty
- English VsTV where V1 is lax and V2 is stressed: *vestigial, despotic*
- V.sTV the lax vowel is not a possible word ending: dε
   Vs.TV the stop is unaspirated but this conflicts with word-initial voiceless stop-V where the stop is aspirated [dεs.paɹɪk]
- Vst.V word-initial vowels have some glottalization, which is missing here
- Compare *asbestos* and *moslemic* where parses are clearer since voiced stop [b] and liquid [l] can occur initially while [zb] and [zl] clusters are absent word-initially

[5] Pritchett's study of expletive infixation in *despotic, mysterious* reports cases of segment splitting where C flanks infix: *des*-fucking-*spotic*

   no-split: Vs-F-tV
   split: V**s**-F-**s**tV
   other: V-F-sTV

	no-split	split	other
i. despotic	12%	**80%**	7%
ii. moslemic	**86%**	0%	14%

- *des*.F.*spotic* matches valid final and initial sequences; *des*.F.*potic* places an unaspirated stop before a stressed vowel and *de*.F.*spotic* ends syllable in a lax nonschwa
- mos.F.lemic is consistent with mo[z].lemic ; #zl is odd and should block the onset parse implicated by the 14%: perhaps they did not voice the [z]

[6]. Malayalam (T. Mohanan 1989 Ling Inq)

- maximizes onsets as evidenced by pa speech disguise

kamala	pa.**ka**.pa.**ma**.pa.**la**	
indu	pa.**yi**.pa.**ndu**	
candran	pa.**ca**.pa.**ndran**	
bandʰanam	pa.**ba**.pa.**ndʰa**.pa.**nam**	'imprisonment'
ammu	pa.?**a**.pa.**mmu**	personal name

cf. Hindi where coda consonant is moraic

tʃandran	pa.**tʃan**.pa.**dran**
bandʰanam	pa.**ban**.pa.**dʰa**.pa.**nam**
ammu	pa.?**am**.pa.**mu**

- Malayalam has a richer inventory of onsets in casual speech:

pḻaawam	'flood'	mḻeeccham	'copper'
kṣama	'patience'	wyasanam	'sorrow'
spaṣṭam	'evident'		
snaanam	'bath'		

- the only final consonants are m and n
- note that Malayalam is consistent with No-Coda >> SSC, *Complex;
- can this grammar with the .CCV preference be found independent of rich initial clusters?

[7] Arrernte (Breen & Pensalfini 1999 LI)

- Lost initial consonants and merged final vowels to schwa, which can now be treated as epenthetic
- Maximal minimal word has VCC underlying structure
- if the medial parse is based on word edges, we predict VC.VCC.VCV with violation of UG Onset and No-Coda
- Breen and Pensalfini present Arrernte as counterexample to the Jakobsonian universal preference for CV syllables and Prince & Smolensky's UG constraints of Onset and No-Coda
- This prediction is supported by several features of the language
- Reduplications

Frequentative:	unt-em	unt-ep-**unt**-em	'run'
	aṭer-em	aṭer-ep-**er**-em	'laugh'
	akemir-em	akemir-ep-**ir**-em	'get up'

Attentuative	ar-em	**ar**-elp-ar-em	'look'
	itir-em	**it**-elp-itir-em	'think'

- Secret Language Rabbit talk (cf. Pig Latin)

emeṇ	eṇem	'plant food'
itirem	irimit	'think'
ulket	etulk	'lizard'

[8] Italian: stop-liquid (TR) as well as s-stop (sT) clusters are possible word-initial sequences just as in English; but they are associated with distinct allomorphs for preceding determiners and contrast in Raddoppiamento sintattico (RS); they also are claimed to be associated with different durations for preceding vowels, which is explained by a tautosyllabic parse for TR and a heterosyllabic parse for sT. (Chierchia 1986, Davis 1990).

ponte	proposito	specchio
il ponte	il proposito	lo specchio
un ponte	un proposito	uno specchio
quel ponte	quel proposito	quello specchio
nessun ponte	nessun proposito	nessuno specchio
'bridge'	'purpose'	'mirror'

pulita	triste	sporca	santa
città	città [tt]riste	città [s]porca	città [ss]anta
città	[pp]ulita	*città [ss]porca	
'clean city'	'sad city'	'filthy city'	'holy city'

[9] This consensus was challenged by Cristie McCrary (2004 UCLA thesis) on the basis of various psycholinguistic and phonetic experiments

syllable parsing experiment
- subjects asked to divide a disyllabic word in speech as well as graphically by drawing a line

Cluster	Task 1 Mean Proportion Tauto (.00 to 1.00 scale)	STDEV	Task 1 Mean rau (-23 to 123 scale)	STDEV	P-value	Predicted
CL	.88	.21	103	30	<.0001	tauto
SC	.59	.38	64	51	.01	hetero
tl	.59	.44	63	63	.06	tauto
SL	.55	.42	57	61	.30	hetero
SN	.41	.45	37	65	.06	hetero
CS	.35	.33	31	45	<.0001	hetero
CN	.35	.35	29	49	<.0001	tauto
CT	.33	.34	26	46	<.0001	hetero
ts	.31	.39	21	55	.0008	hetero
ft	.29	.35	20	49	<.0001	hetero
bn	.28	.40	18	54	<.0001	hetero
GM	.27	.33	18	44	<.0001	tauto
fn	.27	.40	17	58	<.0001	hetero
NC	.08	.21	-10	31	<.0001	hetero
LC	.05	.01	-13	18	<.0001	hetero

Table 2.3 Results of Task 1 in proportions and rau, along with standard deviations and P-values from a one sample two tailed t-test of rau values (hypothesized mean 50). The mean rau values for the clusters in the gray rows of the table are either not significantly different than 50 (e.g. tl, SL, SN) or the differences is marginally significant (e.g. SC). Results located within the same box are not significantly different (p>.01).

- clearest judgments for clusters with sharpest sonority clines: stop-liquid CL (tauto) vs. sonorant-stop NC, LC (hetero)
- intermediate status for sC, sL, sN, and tl
- McCrary sees these results as problematic for the classic analysis but they seem to confirm it since sC and tl are special cases that require additional rules/constraints in Clements & Keyser (1982) and subsequent analyses
- They are problematic if one assumes that subjects behavior is based on surface parsings without regard for the grammatical rules/constraints that produced them

[10] Pisan Italian (McCrary 2004)
- syllable parsing is alleged to be crucial for phonetic duration in previous literature: stressed vowel is longer in open syllable vs. closed
- predicts that rising sonority clusters should be associated with longer vowel duration than level or falling
- not supported by McCrary's (2004) experimental data: CaCCa and CaCa nonse words; 15 Pisan speakers 4 repetitions in frame sentence: 1,851 data points
- V1 duration differences in matched CaTRa vs. CaRTa (e.g. pa.pra vs.par.pa) was not significant for R = liquid for 134 of 15 subjects but was for R = nasal
- Fairly good correlation between V duration and duration of CC cluster regardless of sonority contour
- But the duration of the interludes themselves differs as a function of sonority cline and hetero-vs. tautosyllabic parse and negative association with preceding vowel (p. 223) so

the duration of the interludes cannot be simply the sum of the individual consonants without regard to order, which seems to necessitate some reference to sonority slope and perhaps the syllable difference after all, but...

- duration (ms) of /s,l,r,n/ in three contexts: pá.sa, páp.sa, pás.ta; pá.ra, páp.ra, pár.pa

	s	l	r	n
intervocalic	112	60	33	54
C._V	129	60	50	87
V_.C	120	67	52	80

- contra findings of Farnetani & Kori '86 where coda is longest (a mora?)
- duration of nonnasal stops in three positions

pá.pa	pá.pra, pá.pla	pár.pa, pál.pa
pá.ta	pá.tra, pát.la	pár.ta, pál.ta
pá.ka	pá.kra, pá.kla	pár.ka, pál.ka

- results

	V_V	V_R	VR_
[p]	108	101	146
[t]	101	108	149
[k]	121	113	151

- significantly shorter in V_V and V_R than in VR_
- all are onset positions
- difference due to contexts where contrast with a geminate is possible
- otto, quattro, *quartto
- hence singleton stop is shortened in context where it contrasts with a geminate but not in position where it does not
- this could explain why the cluster in VCRV is shorter than in VRCV rather than location of syllable boundary
- proposed explanation for distribution of geminates
    - o vowel duration ratio is a cue for singleton vs. geminate contrast
    - o hence postvocalic position is optimal for the contrast
    - o nonnasal stops prefer presonorant position: burst and CV formant transitions are most salient to mark the right edge of the stop

- in Latin, An. Greek (Gianninni & Marotta 1989) underlying geminates restricted to intervocalic position--the most optimal position for salient right edge
- Presonorant is optimal context for a geminate vs. singleton distinction since release cues for a stop are most salient here

[11]. Still unexplained is the stress differences:
- antepenultimate stress is possible for CVCVCRV but not CVCVRCV: cf. ánatra 'duck', cáttedra 'teacher's desk', talénto 'talent', paléstra 'gym', pazzésco 'crazy';
- also Latin short mid vowels diphthongized in open syllables: miele, Pietro, dietro, merda, suono, porta,

24.961 Introduction to Phonology
Fall 2014