Workshop: Phonological Typology of Syllable and Word Languages in Theory and Practice
Freiburg Institute for Advanced Studies (FRIAS), Freiburg, 29-31 March 2010

Barış Kabak, *Syllables and emergent segments*

Syllables and emergent segments
Barış Kabak
University of Konstanz

I. Introduction

- Languages that seem to have syllable-based prosodies, rather than word-based prosodies, are those where the syllable is the primary prosodic category that underlies a great majority of phonological regularities.

- My main focus here is on the synchronic, diachronic and psycholinguistic consequences of employing a syllable-based prosody. More specifically, I will explore the effect of being a "syllable" language on (i) the phonological system of the language, in particular on its segmental, morphological and accentual system, and (ii) on the way speech is processed.

- I will first make the observation that the primacy of syllables in languages should not exclude the relevance of other prosodic units. More specifically, I will argue that the Phonological Word constitutes a vital prosodic domain in the languages under exploration here, despite their preferential treatment of the syllable as the basic category.

- Second, I will suggest that the disposition of a "mixed" behavior should not necessarily be seen as evidence or motivation to assume that there is a syllable-language to a word-language drift. Despite the fact that there is overwhelming evidence for syllables being pervasive, rules and regularities making reference to (or being bound to) the Phonological Word (PW) happily co-exist with syllable-based rules. This is compatible with Auer (1993).

- The notion of the PW is vague, and tends to be very language-specific, as opposed to the notion of the syllable. The PW is not necessarily isomorphic to a morphosyntactic word. What looks like a syllable-language phenomenon can easily be interpreted as a PW-related phenomenon, which leads to a paradox. External sandhi processes that mimic syllable-level processes can be viewed as the result of PW-level rules that get extended to higher-level prosodic domains, more specifically the recursive PWs. Furthermore, the syllable languages under investigation often take all kinds of precautions to protect the integrity of bound morphemes. Hence, the rules enhancing the syllable actually protect morphemes. Perhaps, we speak of morpheme-languages?

Phenomena under investigation:
- Emergence of epenthetic vowels and consonants
- Perceptual epenthesis
- The use of harmony and stress as cues for word segmentation
- Emergence and proliferation of bound morphemes and lexical accents

Languages under focus:
- Korean and Turkish, both agglutinative (see Kabak & Plank 2006 for flexive tendencies in these languages).

II. Korean: A syllable- or a word-language?

- The syllable not only lies at the heart of Korean phonological system but also underlies its centuries-old indigenous writing system, Hangul, which is based on morphophonemic syllables.

- The great majority of phonotactic restrictions in fact falls out from the constraints on the syllable structure, which is relatively simple (albeit the on-glide diphthongs complicate the picture a bit).

- The language employs several obligatory phonological processes that refer to syllable positions (e.g., coda neutralization).

(1) Coda neutralization in Korean (Sohn 1999):[1]

- All tense and aspirated stops become plain, and /t, tʰ, s, s', c, cʰ, and h/ are neutralized to [t] in the syllable-final position.[2]

	UR	Surface	
a.	pa**tʰ** # kal-i	ka**t**.k'ari	'field plowing'
b.	o**s** # an	o.**d**an	'upper garment'
c.	u**s** # o**s**	u.**d**ot	'upper garment'
d.	i**c**-ta	i**t**.t'a	'forget'
e.	puə**kʰ** # apʰ	pu.ə.**g**ap	'kitchen front'

- Many processes are there to preserve the syllable structure, which strives to be CV:

(2) Consonant cluster simplification:

a.	nəks	nək	'soul'
b.	pa:lp	pa:p, pa:l	'step on'
c.	ilʰ-ko	il.kʰo	'lose and'

- Phonotactic restrictions and phonological processes at the level of the word (and higher-prosodic domains such as the PPh) exist. Some examples:

  o The absence of word-initial liquids (they can occur elsewhere) is notorious across Altaic languages, also in Korean.

  o [n, l] epenthesis takes place at the beginning of a /i/- or /j/-initial *word crucially after* a consonant-final word.

  o Long vowel shortening

[1] Unless otherwise noted all Korean examples are from Sohn (1999).
[2] Throughout the paper, the left-most examples show the morphophonemic (underlying) representations unless otherwise is indicated. The symbols <.>, <->, <#> refer to a syllable, morpheme, and word boundary, respectively.

Workshop: Phonological Typology of Syllable and Word Languages in Theory and Practice
Freiburg Institute for Advanced Studies (FRIAS), Freiburg, 29-31 March 2010

Barış Kabak, *Syllables and emergent segments*

(3)   [n, l] epenthesis (sensitive to words):

a. cis # iki-ta   ciɲ.ɲi.gi.da   'knead to a mash'
b. k'ocʰ # ilɨm   k'oɲ.ɲi.ɾɨm   'flower name'
c. səul # jək   sə.ul.ljək   'Seoul Station'

- Korean has a segmental rule that is sensitive even to a Phonological Phrase.
  Modern Korean has vowel length distinctions in word initial position, which
  are disappearing as we speak in Korean (see Kabak 2004 for details). Korean
  systematically shortens long vowels in a non-phrase initial syllable.

(4)   Long vowel shortening (sensitive to phrases):
a.   sa:lam   sa:.ɾam   'person'
b.   kʰɨn # sa:lam   kʰɨn sa.ɾam   'tall person'

Assimilation processes:

- There is a considerable number of *strictly obligatory* and highly predictable
  consonantal assimilation processes (except for decoronization, which is
  optional).

- While such drastic assimilatory processes might make Korean appear as a
  word-language, the fact that the rules also apply across words <u>with the same
  strength</u> crucially blurs not only morphological but also word boundaries.

Boundary enhancing	*Boundary weakening*
tensification: (*cap # ci -> cap.c'i 'magazine'*)	**lateralization:** (*ci:n # li -> ci:l.li 'truth'*)
non-application of palatalization: *katʰ-i → ka.cʰi 'together'* *patʰ # ilaŋ → *pacʰiɾaŋ 'plowed row'*	**sibilation:** (*cəc # so -> cə(s).s'o 'milking cow'*)
sai sios (see below)	**nasalization:** (*cip # mun -> cim.mun 'house gate'*)
	decoronization: (*kas # k'ɨn -> kat k'ɨn ~ kak k'ɨn 'hat string'*)

Sai-sios
- Sai-sios (Bindungs-s) between morphemes forming words or in compounds.
  The end result is the intensification of the following lax obstruent, plus other
  assimilatory processes where necessary.

(5)   Sai sios:

a. pi # paŋul {pi-s paŋul}   pit.p'aŋul ∼ pip.p'aŋul   'rain drops'
b. pom # pi {pom-s pi}   pom.p'i   'spring rain'
c. pi # mul   {pi-s # mul}   pin.mul   ∼   pim.mul   'rain water'

- Several minimal pairs exist:

(6)   a. ka:m.k'i   'winding'   (/ka:m-s # ki/)
        b. ka:m.ki   'flu'

- Morpheme and word boundaries are highly blurred in Korean due to the
  <u>uniformity of within- and across-word processes</u>. This is in accordance with
  what one can expect from a prototypical syllable-language. However, one can
  also look at this differently: Korean strives to treat sequences larger than a
  morphosyntactic word as a SINGLE PHONOLOGICAL WORD. Since the rules
  that blur the boundaries are extremely regular, this poses no challenge for the
  learner who has the chance to observe the surface alternations and thus is able
  to construct the underlying morphophonemic representations.

III.   Turkish: A syllable- or a word-language?

- No *obligatory* assimilatory processes across syllables (unlike in Korean). That
  is, Turkish does not blur syllable boundaries. Geminates are tolerated across
  syllables (e.g., *bakkal* 'grocer') as well as across morpheme boundaries (*bit-ti*
  'it finished').

- The language strives to keep its syllables maximally bimoraic (Kabak & Vogel
  2001, Kabak & Revithiadou 2009a), hence <u>underlying</u> (i) consonant clusters of
  a particular type, (ii) geminates, and (iii) long vowels in closed syllables
  cannot surface at the end of a word.

- Such illicit syllable types give rise to processes such as epenthesis , geminate
  simplification, and vowel shortening .

- Syllable repair strategies are, however, blocked when the choice of
  <u>resyllabification</u> is available. This is the case when a following vowel-initial
  element is part of the same Phonological Word (7b), or the recursive
  Phonological Word (PW$_{REC}$) that arise from compound-like constructions (7c),
  and crucially not when it is contained in a syntactic phrase (7d).

(7)   Long vowel shortening:

a. /hara:m/   [har**a**m]$_{PW}$ 'forbidden'
                           *hara:m
b. /hara:m - (y)I/   [har**a**:m-ɨ]$_{PW}$
   forbidden-ACC   *haram-ɨ

Workshop: Phonological Typology of Syllable and Word Languages in Theory and Practice
Freiburg Institute for Advanced Studies (FRIAS), Freiburg, 29-31 March 2010

Barış Kabak, *Syllables and emergent segments*

c. /hara:m   et-mEk/   [hara:m]$_{PW}$ [etmek]PW]$_{PW}$ 'to take the pleasure'
   forbidden LV-INF   *haram etmek

d. /hara:m   et/   [[haram]$_{PW}$ [et]$_{PW}$]$_{PPh}$ 'forbidden meat'
    forbidden meat   *hara:m et

- The optional resyllabification process blocks the word-final devoicing process.

(8) Word-final devoicing:

   a.   /kitab/   ki.tap (*ki.tab)   ki.ta.b-ɨ
                             book   book-Acc

   b.   /kitap # ev-i/   ki.ta**p**.e.vi   ∼   ki.ta.**b**e.vi   'bookstore'
         book#house

- In sum, syllable-repair processes preserve the syllable structure. While they
  are enforced within Phonological Words, they can also apply, albeit with a
  diminished strength, across words within constructions that are arguably
  Recursive-Phonological Words (Kabak & Revithiadou 2009a). So, syllable
  structure rules are actually restricted to the PW, and they can also operate in
  extended PWs.

Vowel harmony and the Phonological Word

- Turkish has regular vowel harmony (palatal and labial harmony), which
  technically operates from one pre-specified vowel to another (so, no need to
  refer to a domain like the PW).

- However, that specified vowels almost always coincide with the word-initial
  syllable make it obvious that we are dealing with a PW-process.

- That word-initial vowels must be pre-specified leads to asymmetries in the
  vocalic properties of word-initial vs. word-non-initial positions (Kabak, in
  press):

  While all 8 vowels are possible in the word-initial position, low rounded
  vowels are not expected in non-initial positions since rounding is not
  supposed to spread onto low vowels.

(9)   a.   soru   *soro   'question'
         b.   kömür   *kömör   'coal'
         c.   soru-lar   *soru-lor   'questions'
         d.   kömür-ler   *kömür-lör   'coals'

- Why would a language have vowel harmony? To neutralize the vocalic
  oppositions within a particular domain, and that IS the PW. As we will see

below, harmony actually serves a cue for word boundaries in online
processing.

The minimal word and category-specific phonology

- A monosyllabic Turkish words must be bi-moraic. There is only one common
  noun that goes against this (*su* "water") but everything else must either be
  lengthened (such as the names for musical notes: *do:*, *re:*, *mi:*, etc.) or extended
  via a glide (*ko-y* 'to put', cf. *ko-du* 'put-past').

- Turkish phonology also cares about whether the element that is subject to the
  rule is a verb or a noun (Kabak & Plank 2007).

- For example, the famous final k-deletion is only specific to nouns
  durak   dura-ın 'bus-stop-Gen'   but   *bırak   *bıra-ır* 'leave-Aor'.

IV.   Vowel hiatus and emergent segments

Mitigation of non-optimal syllables:

- Vowel hiatus resolution: Both Korean and Turkish do not tolerate
  heterosyllabic vowel contact (i.e., V.V). Presumably because syllables must
  have onsets.

Korean:
- Vowel coalescence optionally combines the height attribute of the first vowel
  (V_1) and the frontness of the second (V_2) either morpheme internally or across
  morpheme boundaries to create a new vowel that occupies two vacant vowel
  slots, thereby surfacing as long.

(10)   Vowel coalescence in Korean (from Kang 1999)

   a.   /ai/   [ɛ:]   'child'
   b.   /oi/   [ö:]   'cucumber'
   c.   /po-i-ta/   [pö:da]   'to be seen'
   c.   /onui/   [onü:]   'brother and sister'
   d.   /tʰɨ-i/   [tʰi:]   'to be opened'

Turkish:

- Suffix allomorphy to avoid vowel hiatus (e.g., Lewis 1967, Clements and
  Keyser, 1983: 67-73):

(11)   Hiatus resolution through suffix allomorphy:
          C-final   V-final
   a.   duvar-ɨ   'wall'   b. kapɨ-yɨ   'door'
         wall-Acc   door-Acc
   c.   yap-**ar** 'he/she does'   d. oyna-**r** 'he/she plays'
         do-Aor   play-Aor

Workshop: Phonological Typology of Syllable and Word Languages in Theory and Practice
Freiburg Institute for Advanced Studies (FRIAS), Freiburg, 29-31 March 2010

Barış Kabak, *Syllables and emergent segments*

Conspiracies: Elision→Repair

- Despite the systematic restraints against unwanted syllable types, the language ends up having V.V sequences, which are not optimal. In spoken Turkish, these are resolved in different ways.

1. Borrowings:

(12) a. k**uaf**ör (French: *Coiffeur*)   'hair dresser'
b. n**eo**n   (French: *Néon*)   'neon'
c. mes**a:i** (Arabic: *mesa*[ʔ]*i*)   'work'
d. m**ea**l (Arabic: *me*[ʕ]*al*)   'meaning'

2. Consonantal deletion processes:

(i)   Words containing <ğ> (soft-g):

- Once a phone, the so-called soft-g (<*ğ*>) is now nothing more than an empty consonantal slot in the phonological representation in Standard Turkish (e.g., Clements & Keyser 1983: 67-73, Sezer 1986: 238-239, Kabak 2007).

(13) a.   ağ**ɨ**r   [a**ɨ**r]   'heavy'
b.   yoğurt   [y**ou**rt]   'yogurt'
c.   öğür   [**öü**r]   'to retch'
d.   sɨğar   [s**ɨa**r]   'fits'
e.   ciğer   [c**ie**r]   'liver'

(ii)   /v/, /y/ and /h/ deletion (e.g., Sezer 1986):

(14) <u>Compensatory Lengthening</u>   <u>**No Comp. Lengthening**</u>
a.   /ahmet/ [a:met] 'Ahmet'   e. /tohum/   [to**u**m]   'seed'
b.   /seyret/ [se:ret] 'watch'   f. /deyim/   [de**i**m]   'idiom'
c.   /övmek/ [ö:mek] 'to praise'   h. /över/   [**öe**r]   '(s)he praises'

- The result is V.V. Note that compensatory lengthening is determined by the syllable structure.

(iii)   Intervocalic velar stop deletion in (certain) derived environments:

(15) <u>UR</u>   <u>k→0</u>
a.   /ayak-I/   [aya**ɨ**]   'foot-Acc'
b.   /etek-In/   [ete**in**]   'skirt-Acc'

Vowel assimilation to resolve hiatus (Sezer 1986: 242-243, Kabak 2007a, b):

- Vowel assimilation operates from left to right where the second vowel undergoes complete assimilation to the first.

(16) a.   ağɨr   [a**ɨ**r]   ~   [aar]   *[**ɨɨ**r]   'heavy'
b.   yoğurt [yourt]   ~   [yoort]   *[yuurt]   'yogurt'

c.   öğür [öür]   ~   [öör]   *[üür]   'to retch'

- Not all V.V sequences are subject to vowel assimilation although it is completely predictable (see Kabak 2007 for details). When the phonological environment does not allow coalescence, repair segments emerge.

(17)   Coronal glide epenthesis takes place when the first vowel in the sequence is coronal, or (i.e., when the first one is not coronal) the second vowel is /i/.

	C→0	[e.i]→*[e:]	glide insertion	
a. /şehir/	[şeir]	*[şeer]	[şe**y**ir]	'city'
b. /eğit/	[eit]	*[eet]	[e**y**it]	'educate'
c. /kafein/	--	*[kafeen]	[kafe**y**in]	'caffeine'

- As for other sequences that cannot undergo vowel assimilation, either labial glide epenthesis or velar glide epenthesis depending on the vowel context are other alternative ways to avoid hiatus.

(18)   Labial glide epenthesis:
a. soğan   [soan]   [so**w**an]   'onion'
b. tavuk   [tauk]   [ta**w**uk]   'chicken'
c. doğa   [doa]   [do**w**a]   'nature'
d. dua   -----   [du**w**a:]   'prayer'

(19)   Velar glide epenthesis:
a. sɨğar   [sɨar]   [sɨ**ɰ**ar]   '(it) fits'
b.sɨhhat   [sɨaɫ]   [sɨ**ɰ**aɫ]   'health'
c. saat   ------   [sa**ɰ**at]   'clock, hour'
d. sɨğɨr   ------   [sɨ**ɰ**ɨr]   'cattle'

- Note the similarity between [y, w, ɰ] and [i, u, ɨ]. Crucially, they are in complementary distribution. Furthermore, such repair strategies are actually familiar to us from well-known examples of vowel epenthesis in Turkish.
- The epenthetic vowel in Turkish is always a <u>high</u> vowel, typically emerging in loan words that violate the syllable structure patterns of Turkish (e.g., Yavaş 1980, Kaun 1999).

(20) a. [t**i**ren]   (Fr. *train*)   'train'
b. [t**ɨ**rafik]   (Fr. *traffic*)   'traffic'
c. [b**u**ronşit]   (Fr. *bronchite*)   'bronchitis'
d. [f**ü**lüt]   (Fr. *flute*)   'flute'

- The glide consonants in Turkish are variants of an underlying sonorant, which is used as a default buffer segment.

- The sonorant appears in hiatus resolution contexts and gets assigned either to a syllable nucleus or a non-nucleus slot depending on the syllable structure,

Workshop: Phonological Typology of Syllable and Word Languages in Theory and Practice
Freiburg Institute for Advanced Studies (FRIAS), Freiburg, 29-31 March 2010

Barış Kabak, *Syllables and emergent segments*

where it takes further place specifications depending on the neighboring sounds.

[Sonorant]
([High])

[u] [ɨ] [i] [ü]   [w] [ɰ] [y]
(nucleus positions)   (non-nucleus positions)
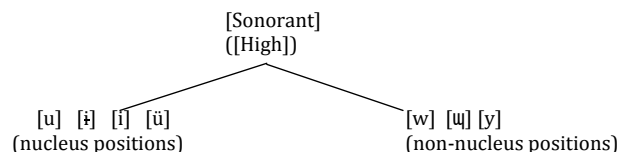
Figure 1: Syllabically-conditioned epenthesis

- All we need to assume then is that both epenthetic vowels and glides are the syllabically conditioned realizations of an unmarked (sonorant) segment used against phonotactic violations (hiatus in this case). Depending on syllable structure demands, it occupies either a non-nucleus position (thereby becoming consonantal) or a nucleus position (thereby becoming vocalic).

ii. Vowel assimilation across words (Kabak 2007b)*:*

- Direction of assimilation is <u>reversed</u>; V_1 assimilates to V_2.

(21) a. ne ayıp!   [naayɨp]   *[neeyɨp]   'how shameful!'
      b. ne okuyor!   [nookuyo(r)] *[neekuyo(r)] 'what does s/he study?'
      c. ne özel!   [nöözel]   *[neezel]   'how special'
      d. anne anne   [annaane]   *[anneene]   'maternal grand mother'
      e. iyi akşamlar   [iyaakşamlar] *[iyiikşamlar] 'good evening'
      f. yirmiüç   [yirmüüç]   *[yirmiiç]   'twenty-three'
      g. ortaokul   [ortookul]   *[ortaakul]   'middle school'
      h. soru işareti   [soriişareti] *[soruuşareti] 'question mark'

- Yet again, the reference to the word is important. This is due to the primacy of word-onsets for lexical access. Languages do not want to manipulate word-initial positions, hence V1, but not V2, changes.

V.   Intolerance for consonant clusters

- Both Turkish and Korean do not permit onset consonant clusters. Speakers often insert an epenthetic vowel into these clusters, as they often come together with loan words typically from Western languages among others.

    (22)   Turkish: *tiren* 'train', *kɨredi* 'credit', *fülüt* 'flute', etc.
    (23)   Korean: *sɨwici* 'switch' *pʰ ɨllajsi*   'flash'

- Strikingly, as said above, both languages allow clusters as well as syllable-structure offending segments (long vowels as well as geminates in the case of

Turkish, see above) in native morpho-phonemic representations, i.e., underlying representations of morphemes. Crucially, they are located at the right-edge, and crucially not at the left-edge, which yet again suggests an asymmetry at the level of the word.

VI.   Perceptual epenthesis:

- Listeners whose native language does not tolerate consonant clusters were shown to hear illusory vowels in consonant clusters (e.g., Dupoux et al. 1999). So, Japanese listeners report hearing a vowel in a word like *ebzo* (as *ebuzo*), and they also cannot distinguish *ebzo* from *ebuzo*.

- But in these studies it is not clear whether epenthesis is due to a purely sequential restriction (i.e., b.z are not attested in Japanese, or due to syllable structure restrictions (i.e., no syllable can end in a [b] in Japanese. Hence *eb.zo* is not a legitimate word). In our previous work (Kabak 2003, Kabak & Idsardi 2007), we tried to tease apart linearly-stated phonotactic restrictions from those that fall out from syllable structure restrictions.

- Korean does not tolerate the heterosyllabic contacts [k.m, l.n, c.m, c.t, j.m, j.t][3] on the surface. But, closer inspection reveals that they do not have the same status. That is, while [k.m, l.n] contain perfectly legitimate consonants in their respective positions (i.e., [k] and [l] are both fine as codas, and [m, n] are fine as onsets), the clusters [c.m, c.t, j.m, j.t] are illicit because they contain a consonant that cannot be in the coda position, namely [c] and [j].

- *Why is this important?*

    1. Which component of L1 phonology constrains the representation and processing of L2 sequences of segments?

    2. Consequences for Linear/ String-based phonotactics (Steriade, 1997; 1999; Blevins, 2003), which state that syllable-based markedness and faithfulness constraints should be stated in string-based terms.

- "What determines positions of contrast or neutralization is the presence/absence and quality (sonorant, obstruent, etc.) of the immediately following or preceding segment" (Blevins, 2003: 383).

- So, there are at least three ways in which a sequence [VC_1C_2V] can be illicit in Korean:
    1. C_1C_2 causes a contact violation (*[k.m], *[l.n])
    2. C_1 causes a coda violation (*[c.] *[.ɾ] *[h.]).
    3. C_2 can't be an onset (*[.l] *[.ŋ]).

---

[3] The Korean [c] and [j] here stand for voiceless and voiced palatal affricates, respectively.

Workshop: Phonological Typology of Syllable and Word Languages in Theory and Practice
Freiburg Institute for Advanced Studies (FRIAS), Freiburg, 29-31 March 2010

Barış Kabak, *Syllables and emergent segments*

- Stimuli:

C_1+C_2 :	Good contact	$[k.t^h], [l.t^h]$
	Bad contact	$*[k.m], *[l.n]$
	Bad on all counts:	$*[c.t], *[j.t], *[c.m], *[j.m]$
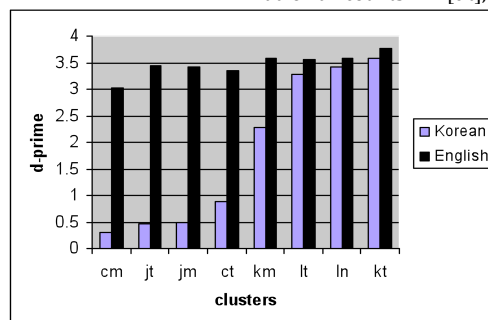
Figure 1: Discriminability scores of heterosyllabic consonant clusters

Bad codas:

$*[p^hac.ma]$	$= [p^ha.c^hi.ma]$	$d'=0.31$
$*[p^hac.t^ha]$	$= [p^ha.c^hi.t^ha]$	$d'=0.89$
$*[p^haj.ma]$	$= [p^ha.ji.ma]$	$d'=0.49$
$*[p^haj.t^ha]$	$= [p^ha.ji.t^ha]$	$d'=0.47$

Bad contacts:

| $*[pak.ma]$ | $≠ [pakuma]$ | $d'=2.28$ |
| $*[pal.na]$ | $≠ [paluna]$ | $d'=3.42$ |

- Results in a nut-shell:
  1. Contact violation is irrelevant.
  2. Bad codas are repaired as onsets.
  3. Perceptual epenthesis emerges to create well-formed syllables.
- Further evidence comes from loan word adaptations:
  Korean does not have [l] in the onset. If a foreign word contains an l-onset, listeners' are faithful to it but the only way to preserve the [l] is to put it in the coda, but this is difficult due to onset maximization. Korean instead doubles the [l] (Kabak 2003).

(24)   Illicit repairs in Korean

	*l.	Onset	Epenthesis	Doubling
L.I. (Linguistic Inquiry)	*ɛ.lay	*ɛl.ay	--	ɛl.lay
igloo	*ig.lu	*igl.u	*igɨl.u	igɨl.lu

- Syllable structure is so pertinacious that surface representations depart too much away from the original, at the cost of having additional segments.

VI.   Syllables and beyond: phonology for word-segmentation

- The culminative nature of accent signals the language user the presence of a lexical word for every (primary) stress-bearing syllable. When coupled with the edge-demarcation function of stress in a given language, this should then signal the language user the existence of word onsets or offsets relative to the position of the stressed syllable in the speech string.
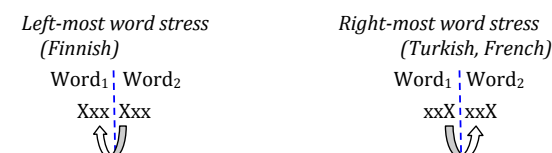
Left-most word stress (Finnish)	Right-most word stress (Turkish, French)
$Word_1	Word_2$	$Word_1	Word_2$
Xxx Xxx	xxX xxX

Figure 3: Demarcative function of edge-most stress (from Kabak, Maniwa & Kazanina 2010)

- Vowel harmony regulations can also help in demarcating individual words since they dictate which vowel features are (dis)allowed to co-occur within a word. Exploiting this abstract knowledge, the language user should then be able to postulate a word boundary whenever a sequence of two adjacent vowels is disharmonic.

	Harmony-match	Harmony-mismatch
Stress-2	goLUshopaVO	**gøLYshø**paVO
Stress-3	goluSHOpaVO	**gølySHØ**paVO

Table 1. A full sample set of conditions used in Kabak et al. 2010. The target word is *paVO*. The stressed syllables are in capitals. Syllables with front vowels are in bold.

- Hypotheses:
  a. Stress should signal a word boundary immediately after the stressed syllable (i.e., in stress-3 condition), thereby facilitating the detection of a target nonword that follows such a boundary in both languages.
  b. Target-word detection should yield faster RTs and/or lower error rates in the harmony-mismatch condition, only in Turkish.

- The results strongly supported this hypothesis. Whenever the target non-word was immediately preceded by a stressed syllable, its detection was faster than when it was not in both languages.

Workshop: Phonological Typology of Syllable and Word Languages in Theory and Practice
Freiburg Institute for Advanced Studies (FRIAS), Freiburg, 29-31 March 2010

Barış Kabak, *Syllables and emergent segments*

- Vowel disharmony also robustly facilitated target detection only in Turkish: targets were detected significantly more accurately in the harmony-mismatch conditions than in the harmony-match conditions.

- Predictable edge-most stress can serve as a word-boundary inducer is that it was found to operate both as a regressive, as Suomi et al. (1997) and Vroomen et al. (1998) found, and as progressive cue, as in Turkish and French. The fact that the effect was robustly present only for the Turkish listeners suggests that language-specific phonological regularities that underlie segmental co-occurrence patterns can influence speech segmentation.

VII. Influence of syllable-structure on morphological and prosodic representations

- In a predictable word-final stress language like Turkish (Kabak & Vogel 2001, in print), accent-bearing morphemes (lexical accents) are not expected. However, there are several morphemes that deviate from the regular word-final stress pattern (stressed vowels are capitalized).

(25) Right-most regular stress:
    a. kitAp 'book'
    b. tanı-dık-lar-ım-dAn 'from my acquaintences'

    Pre-accenting morphemes:
    c. yAp-ma-dı '(s)he did not do'
    d. annE-ymiş 'she is a mother' (hearsay)

    Locally accented morphemes:
    e. gel-Iyor '(s)he is coming'
    f. yap-Amaz '(s)he cannot do'

- In a regular right-most language like Turkish, the emergence of lexical accents has been attributed to the grammaticalization of compound-type structures. Crucially, the conflicting directionality of compound-stress assignment in comparison to word-stress assignment determines whether an edge-most language will develop lexical accents (see Kabak & Revithiadou 2009b).

- In Turkish, compound stress is as follows: The stress of the left-most phonological word is promoted:

(26) a. [kitAp] [ev-I] 'bookstore'
      book house-CmPM
    b. [kurU] [fasulyE] 'dried beans'
      dry bean

- Arguably, lexical accents in Turkish came from compositional constructions. Essentially, the second word was morphologized and became dependent on the first one, causing the stress to appear immediately before it, leading to the emergence of pre-accenting morphemes. Notice also that the bound form undergoes vowel harmony with the host it attaches to.

(27) a. [adAm] # [i-sE] > adAm=sa 'as for the man'
      [melEk] # [i-sE] > melEk=se 'as for the angel'
    b. [adAm] # [ilE] > adAm=la 'with the man'
      [melEk] # [i-lE] > melEk=le 'with the angel'
    c. [parA] # [i-sE] > parA=ysa 'as for money'
      [dedE] # [i-sE] > dedE=yse 'as for the grandfather'
    d. [parA] # [ilE] > parA=yla 'with money'
      [dedE] # [i-lE] > dede=yle 'with the grandfather'

- It looks like a whole syllable is deleted in (a, b) when the second member of these constructions is grammaticalized. However, this is not so when we consider (c, d): the segmental effect of the deleted vowel is retained.

- The [y] appearing with the bound forms in (c, d) has nothing to do with the so-called buffer consonants that emerge in hiatus resolution processes discussed above. It emerges for no syllable repair reason (i.e., [y] should be inserted when there is a vowel hiatus context. Just –*sa/-se* and –*la/-le* would have been sufficient.

- Here, it could be suggested that the bound form, which must not be identical to the full form, is maximally faithful to the featural content of the full form unless a syllable structure violation is incurred.

- Accordingly, =*yle* and =*yse* are maximally similar to *ile* and *ise*, respectively. Hence, the deletion of the glide in post-consonantal contexts finds a plausible explanation: the presence of *y* creates a complex onset following consonants, leading to a syllable structure violation (**melek.yle*); thus, it must be deleted (see Kabak 2006 for details).

- In certain cases, the final suffix of the first word and the following element were reanalyzed as one morphological unit, these typically carry a lexical accent on their initial syllable.

(28) a. [düş-E] # [yaz] > düş-<u>Eyaz</u>
    b. [al-A] # [yorur] > al-<u>Iyor</u>
    c. [gel-I] # [ver] > gel-<u>Iver</u>

- All locally accented morphemes are <u>disyllabic</u> in Turkish. The reason is that the syllables did not elide or fused when reanalysis took place.

Workshop: Phonological Typology of Syllable and Word Languages in Theory and Practice
Freiburg Institute for Advanced Studies (FRIAS), Freiburg, 29-31 March 2010

Barış Kabak, *Syllables and emergent segments*

- Having fixed (edge-most) word stress, which is a property of syllable-languages, has consequences for the type of lexical accents that can emerge in the course of grammaticalization.

VIII. Conclusions

- Syllables are pertinacious in Korean and Turkish. They shape morphophonological representations and are at the heart of processing.

- These languages also make use of regularities that refer to the word, which is only natural if one assumes the prosodic hierarchy and considers the non-isomorphism between the Phonological Word and morphosyntax. That is, Phonological Words can be bigger than morphosyntactic words. Hence, what appears to be external sandhi phenomenon is actually a PW-level process that gets extended.

- Vowel harmony is truly a PW-related process although it is correlated with being a syllable-language.

- We have a paradox here, but there is a resolution: the typology should be based on syllables, not words.

- More cross-linguistic experimental research is needed to understand the psycholinguistic consequences of using syllables as a primary unit in languages for the nature of phonological processing and representations.

References

Auer, Peter. 1993. *Is a rhythm-based typology possible? A study of the role of prosody in phonological typology*. Arbeitspapier 21. University of Konstanz.

Clements, George N. & Engin Sezer. 1982. Vowel and consonant disharmony in Turkish. In: Harry van der Hulst and Norval Smith (eds.), *The Structure of Phonological Representations*, Part 2, 213-255. Dordrecht: Foris.

Blevins, Juliette. 2003. The independent nature of phonotactic constraints: an alternative to syllable-based approaches. In: Féry Caroline. & Ruben van de Weijer (eds.), *The syllable in Optimality Theory*, 375-403. Cambridge: Cambridge University Press.

Dupoux, E.; Kakehi, K.; Hirose, Y.; Pallier, C.; Mehler, J. 1999. Epenthetic vowels in Japanese: A perceptual illusion? *Journal of Experimental Psychology: Human Perception and Performance* 25: 1568-1578.

Kabak, Barış & Irene Vogel. 2001. The Phonological Word and Stress Assignment in Turkish. *Phonology* 18: 315-360.

Kabak, Barış. 2003. *The Perceptual processing of second language consonant clusters*. Ph.D. Dissertation. University of Delaware. UMI Dissertation Publishing.

Kabak, Barış & William Idsardi. 2003. Syllabically Conditioned Perceptual Epenthesis. In: Nowak, P. et al. (eds.). *Proceedings of the Berkeley Linguistics Society 29.* pp. 233-24

Kabak, Barış. 2004. Acquiring phonology is not acquiring inventories but contrasts: The loss of Turkic and Korean primary long vowels. *Linguistic Typology* 8: 351-368.

Kabak, Barış. 2006. An obstacle to the morphologization of postpositions. *Studies in Language* 30: 33-68.

Kabak, Barış & Frans Plank. 2006. Where flexion encroaches on agglutination in Turkish and Korean. In: Yong-Kun Ko et al. (eds.) *Whither Morphology in the New Millennium*?, Morphology Monograph Series 1. pp. 123-152. Seoul, Korea: Pagijong Press.

Kabak, Barış. 2007a. Hiatus resolution in Turkish: an underspecification account. *Lingua* 117:1378-1411.

Kabak, Barış. 2007b. Vowel assimilation across words in Turkish. *Turkic Languages* 11: 181-195.

Kabak, Barış & William Idsardi. 2007. Perceptual distortions in the adaptation of English consonant clusters: Syllable structure or consonantal contact constraints? *Language and Speech* 50: 23-52.

Kabak, Barış & Frans Plank. 2007. Phonology and word class distinctions. Paper presented at the 7th Biennial Meeting of the Association for Linguistic Typology, Paris, France.

Kabak, Barış & Anthi Revithiadou. 2009-a. An interface approach to prosodic word recursivity. In: J. Grijzenhout & B. Kabak (eds.) *Phonological domains: universals and deviations*, 105-133. Berlin: Mouton de Gruyter.

Kabak, Barış & Anthi Revithiadou. 2009-b. From edgemost to lexical stress: Diachronic paths, typology and representation. *The Linguistic Review* 26, 1-36.

Kabak, Barış, Kazumi Maniwa, & Nina Kazanina. 2010. Listeners use vowel harmony and word-final stress to spot nonsense words: A study of Turkish and French. *Laboratory Phonology* 1: 207-224.

Kabak, Barış & Irene Vogel. in press. Exceptions to stress and harmony: cophonologies or prespecification? In: Horst Simon & Heike Wiese (eds.) *Expecting the unexpected: exceptions in grammar.* Mouton de Gruyter.

Kabak, Barış. in press. Turkish vowel harmony. In Marc van Oostendorp, Colin Ewen, Beth Hume, & Keren Rice (eds.), *The Blackwell companion to phonology*. Wiley-Blackwell.

Kang, Ongmi. 1999. 'A correspondence analysis on hiatus resolution in Korean', *Studies in Phonetics, Phonology and Morphology 5*.

Kaun, Abigail. 1999. 'Epenthesis-driven harmony in Turkish', *Proceedings of the Berkeley Linguistics Society* 25, 95-106.

Lewis, G. Lewis. 1967. *Turkish Grammar*. Oxford: Oxford University Press.

Sezer, Engin. 1986. An autosegmental analysis of compensatory lengthening in Turkish. In: W. L. Wetzels & E. Sezer (eds.) *Studies in Compensatory Lengthening*, 227-250. Dordrecht: Foris.

Sohn, Ho-Min. 1999. *The Korean language*. Cambridge: Cambridge University Press.

Steriade, Donca. 1999. Alternatives to syllabic analysis of consonantal phonotactics. In: Osamu Fujimura, Brian Joseph and Bohumil Palek (eds.) *Proceedings of the 2nd Linguistics and Phonetics Conference*, 205-245. Prague: Charles University Press.

Suomi, K., McQueen, J. M., & Cutler, A. 1997. Vowel harmony and speech segmentation in Finnish. *Journal of Memory & Language* 36: 422-444.

Yavaş, Mehmet S. 1980. *Borrowing and its implications for Turkish phonology*. Unpublished Ph.D. Dissertation. University of Kansas.

Vroomen, J., Tuomainen, J., & de Gelder, B. 1998. The roles of word stress and vowel harmony in speech segmentation. *Journal of Memory and Language* 38: 133-149.

Abbreviations

Aor	Aorist	Pr	Person (e.g., 3P= 3rd person)
Abl	Ablative	Pl	Plural
Acc	Accusative		
CmP	Compound Marker		
Fr	French		
Gen	Genitive		
Loc	Locative		

Address for correspondence:

Baris.Kabak@uni-konstanz.de

Universität Konstanz
FB Sprachwissenschaft
Fach 180
D-78457 Konstanz
Germany