Lexical phonology

Marc van Oostendorp

December 6, 2005

Background

- Until now, we have presented phonological theory as if it is a monolithic unit.
- However, there is evidence that phonology consists of at least two layers; that in many natural languages there is a difference between phonology as it is applied within words (lexical phonology) and phonology as it is applied across words (postlexical phonology)

1 Two layers of phonology

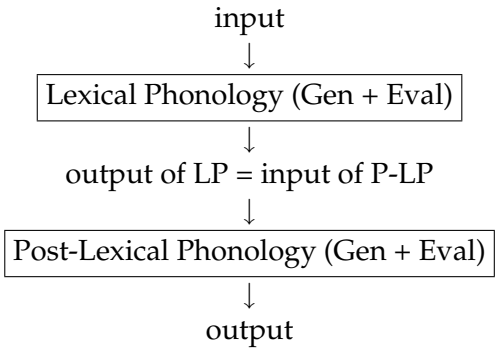
Consider the following facts of Dutch phonology:

(1)	/ɪk hɛb ət/	[ɪk.hɛ.pət]	'I have it'
	/hɛi hɑd ət/	[hɛi.hɑ.tət]	'He had it'
	/ɪk hɛb ər/	[ik.hɛ.pər]	'I have her'

The question which raises about these examples is: why are the final obstruents of *heb* and *had* devoiced? We know by now that Dutch has a process of syllable-final devoicing: the feature [voice] is not allowed to appear in the syllable coda. But the strange thing is that in cases such as this, the obstruents in question do not appear in the coda: they are in the onset of the next syllable, since Dutch syllables cannot begin with a schwa.

The point about these forms of course is that the obstruents are at the end of a syllable if we syllabify these words disregarding their syntactic context. One of the ideas of the theory of Lexical Phonology is that this idea is essentially correct: there are (at least) two phonologies — each an Optimality Theoretic system — which are serialised: first we apply phonology to words (lexical phonology) and the output of this is then, after the operation of syntax, applied to phrases (postlexical phonology):

(2)



In the Dutch case at hand, this works as follows. The input form is { /hɛb/, /ət/, /ɪk/ } (the order is arbitrary). The output of the Lexical Phonology will be { [hɛp], [ət], [ɪk] }, with still an arbitrary order. This will also be the input of the Post-Lexical component, except that the syntax will determine word-order at some point. The output of this component will be [ɪk.hɛ.pət]. The [p] will be in the onset, but there is no reason why it should turn into a [b]; as far as the Post-Lexical phonology is concerned, [p] is completely faithful.

Given this serialisation, there are still two possibilities: the Lexical and the Postlexical grammar could be internally exactly the same (be composed of the same ranking of the same constraints) or they could be completely different. The empirical facts seem to point in the second direction in many languages. For instance, in French, the surface syllable structure seems much more complex than what is allowed at the word level. Although there are no French words which start with the cluster *[dvr], we do find such clusters postlexically:

(3) *Henri devrait partir* 'H. would have to leave' [ãridvrɛpartir]

Similarly, in Dutch we may find 'onset' clusters such as [tf] or [kx], but only at the postlexical level (*'t Valt niet mee* 'It's hard' [tfɑltnitme], *'k Geloof er niks van* 'I don't believe a word of it' [kxələofərnɪksfɑn]).

It thus seems that the restrictions on syllable structure at the postlexical level are more relaxed than at the lexical level: whatever the constraints against complex clusters are, they are more highly ranked lexically than they are postlexically. For example:

(4) a. Lexical ranking: COMPLEX≫FAITHFULNESS
b. Postlexical ranking: FAITHFULNESS≫COMPLEX

In the course of time, phonologists have compiled a list of general differences between lexical and postlexical processes (Kenstowicz, 1994):

(5) a. Lexical processes are usually sensitive to lexical information (lexical exceptions etc.), postlexical are not sensitive to such information.

b. Lexical processes change one phoneme into another, postlexical rules are typically about allophonic changes.
c. The phonetic motivation for postlexical processes is often completely transparent, whereas this is not the case for lexical rules.

By way of an illustration, two processes in English phonology are sometimes used: the lexical process of trisyllabic shortening, and the postlexical process of flapping.

Trisyllabic shortening is the name for a process which shortens a vowel when it is followed by two other vowels:

(6) a. *div*/aj/*n* + *ity* → *div*[ĭ]*nity*
b. *gr*/eː/*de* + *ual* → *gr*[ĕ]*dual*
c. *cl*/ij/*r* + *ify* → *cl*[ă]*rify*

Note that the quality of the vowel also changes, but that is something which is usually abstracted away from. We can describe this by the following constraint:

(7) TRISYLLABIC SHORTENING (TSL): A long vowel cannot be followed by two short vowels.

According to all three criteria in (5), TSL is a lexical process:

a. It is subject to lexical factors. There are many words to which is does not apply, such as *nightingale* and *ivory*. In particular, TSL can only be seen at work with certain Latinate suffixes, such as *-ity*, *-ual* and *-ify*. Other suffixes do not trigger the process at all:

   (8) *br*[eː]*very*, *m*[aj]*tily*, *p*[aj]*rating*

   And even if a suffix will normally trigger TSL, it might not do so in certain individual cases: *ob*[iː]*sity* does not turn into **ob*[ĕ]*sity*.
b. Any change for which TSL is responsible will turn one phoneme of the language into another phoneme. All the sounds we have mentioned so far {aj, iː, eː, ă, ăe, ĕ,...} represent (distinct) phonemes of the language. It is said that lexical processes are **structure preserving**: they never create sounds which do not already exist underlyingly.
c. It is not very easy to describe the process in a 'natural' way. As a matter of course, the formulation of TSL in (7) is hardly satisfactory; we would like to reduce it to what we already know about trochaic feet. But has proved to be not very easy to find a good formulation of the process in question in these terms. It is also not easy to find evidence for TSL in other languages of the world.

TSL contrasts in all these respects from so-called flapping, whereby an intervocalic coronal stop /t/ or /d/ turns into a sonorant flap — which we will write as [D] — between a stressed vowel and a following unstressed vowel:

(9) a. *a/t/om* → *a[D]om* (cf. *a/t/ómic*)
b. *mee/t/* + *ing* → *mee[D]ing*
c. *what/t/* + *is wrong* → *wha[D] is wrong*

We can formulate the following constraint:

(10) FLAP: (Coronal) sonorants are preferred over (coronal) stops in the middle of a foot.

Let us check how this process fares with respect to the three criteria in (5):

a. The process is not subject to any lexical factors: all relevant underlying coronal stops will flap if they occur in the right context. As the examples in (9) show, the process occurs both within morphemes as well as across morpheme and even word boundaries. There are no lexical exceptions.
b. The output [D] is not an independent phoneme of English; there are no minimal pairs distinguishing /t/ (or /d/) from a hypothetical /D/. We actually only find this sound as a result of flapping, and untrained speakers are typically unaware of the fact that they say a different consonant in *atom* than in *atomic*.
c. Finally, FLAP has a natural phonetic explanation: onsets of stressless syllables tend to 'assimilate' in sonority to the surrounding vowels. This type of process — **lenition** — is also something which is well attested in the world's languages.

It thus looks as if the Lexical Phonology is a different type of grammar from the Post-Lexical Phonology; while the former is much closer to the lexicon — hence the sensitivity to exceptions and to morphological structure, and the structure preservation — th latter is much closer to the phonetics — hence the naturalness. Especially for the extreme cases of both, some scholars have argued that they are 'not real phonology':

- For lexical phenomena, some say that they are really lexical, hence listed in the lexicon, and not the result of some grammar. A speaker of English has *clărity* in her head, and she does not produce this on line from *clear* and *-ity*. A less favourable resukt of this line of thinking is that many regularities cannot be accounted for.
- For postlexical phenomena, some say that they are really phonetic: flapping is just an automatic consequence of the way the human speech organs work. A potential problem with this is that it cannot account for the fact that languages may differ (also) with respect to post-lexical

phenomena, and this is not something we want to account for in 'universal' phonetics.

It is a consequence of the organisation of the grammar in (2) that lexical processes should precede the post-lexical processes. The English phenomena we have just considered do not provide us with a lot of evidence about this, but in cases where there is evidence of ordering, it basically always goes in the right direction. For instance, two things happen in the following Polish examples:

(11)

bup	bob-u	'bean'
xut	xod-u	'pace'
kot	kot-a	'cat'
vus	voz-u	'cart'

In the first place, Polish is subject to syllable-final devoicing, a process we have already seen operative in many other languages. This process is thus 'natural', it is also exceptionless in Polish and it replaces phonemes by other phonemes. It thus is a paradigm case of a postlexical process.

In the second place, there is a process raising underlying /o/ to [u] when it occurs before a word-final voiced consonant, explaining the alternation in e.g. *bup* from underlying /bob/. A form such as *kot* illustrates that raising never happens if the form underlyingly has a voiceless obstruent.

Raising is a textbook case of a lexical process. It has lexical exceptions: for instance the words *mob* [mop] and *snob* [snop] do not undergo it. It talks about two phonemes, [u] and [o]. Finally, there seems no phonetic motivation for this type of raising in this particular context.

We have a clear indication that these processes must be ordered: raising does not see the effect of devoicing, since it still distinguishes between /xod/ and /kot/. In the first place, this means that we cannot do with only one grammar, just as in the case of Dutch. This only works well if we order raising before devoicing:

(12) a.

	/xod/	/kot/
raising	xud	n.a.
devoicing	xut	n.a.
	[xut]	[kot]

b.

	/xod/	/kot/
devoicing	xot	n.a.
raising	d.n.a.	n.a.
	*[xot]	[kot]

The comparison of Dutch and Polish teaches us, by the way, that a process (final devoicing) which is lexical in one language may be postlexical in another one. Usually, it is assumed that phonological processes start their life

in natural language as postlexical processes, only to turn lexical in the course of time.

2 Cyclicity

We can thus see that phonology is not monolithic. There are at least two phonological modules. In the architecture of the grammar, it would be logical to say that one precedes the syntactic component, whereas the other follows it. The former thus interacts with morphology, and the latter with the output of syntax.

We will now concentrate on the lexical phonology. Until now, we have concentrated in this course on words which consist of one or two morphemes. But what happens if there are more than two morphemes in the input, viz. a stem plus more than one affix? In principle, there are two possibilities: we add all the affixes at the same time, or we first add one affix, then we apply phonology, and then we add another affix. It is one of the basic tenets of so-called cyclic phonology — usually incorporated into Lexical Phonology — that the latter is the case. The model is called cyclic, because we go in a circle: we add a suffix, then we apply phonology, then we add another suffix, etc. In this class we will assume that the phonology is the same constraint ranking every time.

A classical example is the difference between the English words *condensation* and *compensation*. (The relevance of the example is not uncontested, but we will use it here for illustrative purposes.) These words have virtually the same segmental makeup, but there stress pattern is different, and so is schwa reduction as a consequence of this: *condɛ̀nsátion* versus *còmpənsátion*. According to this analysis, the reason for this is that the former noun is related to the verb *condɛ́nse*, whereas the latter is related to *còmpənsáte*. These relations are reflected in the stress structure of the verbs.

The cyclic account of this would run along the following lines. We first derive the stress structure of the verbs, then we attach the suffixes and — on the next cycle — we get the stress structure of the nouns.

(13)	underlying	*compensate*	*condense*	
		↓	↓	stress assignment
		còmpənsáte	*condɛ́nse*	
		↓	↓	nominalisation
		còmpənsáte+ion	*condɛ́nse+ation*	
		↓	↓	stress assignment
	surface	*còmpənsátion*	*condɛ̀nsátion*	

The working of the cycle is one of the key components of traditional generative phonology. Within generative syntax (minimalism) it still survives in the

form of derivational phases; it may be considered as one of Chomsky's most important contributions to our insight into linguistics.

A more complicated — and realistic — example comes from Palestinian Arabic (Brame, 1974; Kager, 1999; Kiparsky, 2000). This language has a process deleting unstressed high front vowels /i/ in all syllables of the word except for the last one:

(14) a. /fihim/ 'to understand' (verb stem)
b. /fihim/+∅ → [fíhim] 'he understood'
c. /fihim/+/na/ → [fhímna] 'we understood'
d. /fihim/+/u/ → [fíhmu] 'they understood'

Now consider the following form:

(15) /fihim/+∅ + /na/ → [fihímna], *[fhímna] 'he understood us'

/i/ deletion does not take place here, even though its phonological conditioning is met. The reason for this is that structures with object clitics are built on the basis of structures which occur as independent words, and we have the following principle governing derivations:

(16) *Natural Bracketing Hypothesis*. A substring ψ of a string ϕ is a domain of cyclic rule application in phonology only if it shows up elsewhere as an independent word sequence which enters compositionally into the determination of the meaning of ϕ.

Since *fíhim* 'he understood' occurs independently in Palestinian Arabic (e.g. in the sentence *fíhim il-waláad* 'he understood the boy'), and since it enters compositionally into the determination of [fihímna] — we have to calculate what the former means in order to understand what the latter means —, we also phonologically determine the structure of 'he understood' in order to get the structure of the whole form. On the other hand, *fíhim* 'he understood' itself derives from the verbal stem *fihim*, but this is not a separately pronounceable word, hence it cannot influence the structure of the stem in any way.

Note that the example looks very much like the English case discussed above, except that in English we were dealing with vowel reduction rather than vowel deletion. There is another difference as well: for the English case we could assume that it would be the stress which is transferred from the base to the extended form and that there would be some secondary stress on the [ɛ] in *condensation*, blocking its reduction. Yet this is not true for the Palestinian dialect of Arabic. There is no trace of secondary stress in the cliticised form. This becomes apparent in parallel forms with other vowels than /i/.

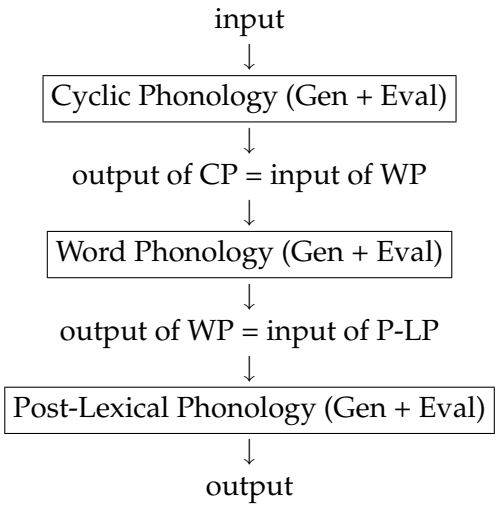
These vowels are not deleted, and words like *ḍarábna* are ambiguous between the meaning 'he hit' and 'he hit us'.

A basic assumption of the model is that cyclic phonology is always lexical (because cyclicity involves sensitivity to morphological structure). Is the reverse also the case? The attentive reader who studies the following Dutch facts will notice that this cannot be the case:

(17) a. *Derivation*: *he/*b/+*erd* → [hɛbərt] ('greedy person')
b. *Inflection*: *he/*b/+*en* → [hɛbən] ('have' PLUR)

We have seen above that devoicing crucially applies lexically in Dutch. These facts show, on the other hand, that it cannot apply cyclically, otherwise it would apply in the stem *heb*. The usual way of thinking about this is that the lexical phonology actually itself consists of two subphonologies: a cyclic component first, which interacts with the morphology in the way we have seen for English and Palestinian Arabic, and a 'word' phonology after this, which applies to whole words before they are then inserted into the syntax. All in all, the standard model of Lexical Phonology thus looks as follows:

(18)



Bibliography

Brame, Michael (1974). 'The cycle in phonology: stress in Palestinian, Maltese and Spanish'. *Linguistic Inquiry*, 5: 39–60.
Kager, René (1999). 'Surface opacity of metrical structure in Optimality Theory'. In: Hermans, Ben & Marc van Oostendorp (eds.), *The Derivational Residue in Phonological Optimality Theory*. Amsterdam: John Benjamins.
Kenstowicz, Michael (1994). *Phonology in Generative Grammar*. Oxford, UK: Blackwell.
Kiparsky, Paul (2000). 'Opacity and Cyclicity'. *The Linguistic Review*, 17: 351–367.

Exercise

Consider the English alternation *electric-electricity*. Is this a lexical or a post-lexical process? Provide arguments.