On the notion *word* and its role as a phonological constituent

Katalin Balogné Bérces

One of the oldest problems of linguistics is the existence of phenomena which are so basic and elementary, and thus unavoidable, in language, and still, whose proper definition, desirable in a scientific study of human communication, is hard to find. One such notorious example is the notion of the syllable, which, being real to native speakers only, and lacking a single established and generally accepted definition both phonetically and phonologically, has been abandoned in phonological theories several times (the most well-known cases include early Generative Phonology launched in Chomsky and Halle (1968) (frequently referred to as SPE), and a more recent research programme called Government Phonology (Kaye et al. 1985, Kaye et al. 1990, Lowenstamm 1996, Szigetvári 1999)).

The story of the word is similar, so much so that the simplest and shortest definition it can be given makes reference to its "universal intuitive recognition by native-speakers, in both spoken and written language" (Crystal 2003: 500). The trouble is only doubled by the usual problem with native intuition, namely, the fact that native speakers of languages of different structural types have different conceptions of what a word is and therefore make different judgements. A string like *Qasu-iir-sar-vig-ssar-si-ngit-luinarnar-puq* (tired-not-cause-place-suitable-find-not-completely-someone-3/sg 'Someone did not find a completely suitable resting place') may be considered as a word by a speaker of Inuktitut (one of the Eskimo languages of North America) but as a sentence by an Englishman.

Even within the same language there are debatable cases. Compounds like *rainbow*, *pickpocket* or *lifestyle* always pose problems: how many words are they? Are compound words single words, as their name suggests? Or are they combinations of two (or more)? The English native speaker’s indecision is also reflected by the indecision of the spelling. Most compounds have two or three alternative spelt forms (e.g., *lifestyle* can be *life style*, *life-style*, or *lifestyle*). Clearly, this is a case of one word (in one sense of the word *word*) containing two words (in a slightly different sense of it). The mirror image of this problem is presented by "words" like *write*, *writes*, *writing*, *wrote*, *written*, which seem to contain the same underlying "word". Clearly, all the questions arising in connection with the definition of the word stem from the non-homogeneity of the notion: it is a linguistic commonplace that there exists a variety of the senses of the word *word*.

One quite obvious sense is the orthographic word: that object in spelling which stretches from one space to another. Although the importance of spelling as a reflection of the language is often over-emphasised in literate cultures, it can be conceived of as an indirect indicator of native intuitions about wordhood (recall: it is this native intuition that defines the notion) as well as an indicator of the history of the language. In the case of English compounds, for instance, it may reflect the degree to which the coinage has become integrated into the system (e.g., professionals frequently using a given compound will tend to start spelling it as one single word whereas laymen may spell the terms separately or hyphenated). This meaning of *word* comes closest to it as a physical object.

The most abstract meaning, on the other hand, is the common underlier of word forms like *write*, *writes*, *writing*, *wrote*, *written* above, the so-called lexeme, accompanied by another sense of the word, referring to the word forms themselves. In fact, the word (= word form) *write* corresponds to several different words (called grammatical or morphosyntactic words): the non-finite infinitive plus a number of finite forms (the present tense forms other than third person singular). Therefore, depending on the meaning of *word*, the list *write*, *writes*, *writing*, *wrote*, *written* may be regarded as representing one, five, or ten words. Out of
these meanings, the word form is the least abstract, and as such, it corresponds to the orthographic word.¹

Last but not least, there is another physical sense of the word word, the pronounced word-form, i.e., the phonological or prosodic word (henceforth, p-word), as opposed to the other, non-phonological meanings mentioned above, which are sometimes subsumed (by phonologists) under the heading of grammatical word (further complicating the picture). This phonological sense of the word word is the one the rest of the paper will concentrate on, and, as it will become clear below, it is not a concept without difficulties of description, either. Namely, it will be pointed out that the p-word is not simply the pronounced (i.e., phonetic) version of the grammatical word.

The main aspect of the p-word that I would like to show is its non-isomorphy with the grammatical word, and the problem of compounds, referred to above briefly, will come in handy as an illustration. What is special about compounds is that, once formed, they behave as a single grammatical unit both morphologically (the terms do not take affixes separately, e.g., the plural of lifestyle is lifestyles and not *livesstyles, that is, the morphological structure of the compound is [[life][style]], where morphological constituents are given in square brackets; pickpockets does not refer to a person ‘picking’ several pockets, but several persons from that walk of life) and syntactically (the terms are not manipulated by the rules of sentence structure separately but the whole compound is inserted as the terminal node of syntactic trees). It is phonology that treats them as separate entities (at least in English and Hungarian, and languages like English and Hungarian). In English, this can be illustrated by the stress pattern of compounds: usually, the first term carries primary stress but the other term(s) also preserve(s) the position of the original main stress in the form of what some call tertiary stress. Thus, the only syllables of the separate words life and style carry primary stress (conforming to the requirement in English that each non-function word contain one (and only one) primary stressed syllable), whereas in lifestyle only the first term retains it and the second one is reduced to tertiary. The point to be noted here is that the second syllable of lifestyle does not reduce completely, i.e., to a schwa sound, which is the usual state of affairs in words other than compounds, cf. shyster.

The same point is illustrated by the historical change of a compound into a non-compound, accompanied by several segmental changes due to the pressure to obey the phonotactic constraints (i.e., constraints on the combination of sound segments) of non-compounds. The word shepherd, for instance, originally a compound from sheep and herd (as its spelling indicates), is treated as a single phonological word in Present-day English, and therefore has lost the stress on the second syllable (which now contains a schwa) as well as the sound /h/ (which is always lost at the beginning of completely unstressed word-internal syllables, cf. vehicle and vehicular; cf. also the two pronunciations of forehead). Other examples of former compounds now containing a reduced vowel in the original second term include cupboard (where the internal consonant cluster has also become simplified owing to the intolerance of monomorphemic /pb/ sequences), handsome, breakfast, postman, orchard and vineyard (where the shortening of the stressed vowel from the /au/ of vine to the /i/ of vineyard, the same phenomenon as the one witnessed in sheep vs. shepherd, is another characteristic of non-compound words, cf. five – fifth, creed - credible). The process whereby a compound loses (some of) its internal morphological structure is usually accompanied by semantic drift – that is, the meaning of the resulting non-compound can no more be reproduced from the meanings of its constituent parts (cf., e.g., the present-day meaning of cupboard). How this happens is illuminated by Kristó (this volume).

¹ For a more detailed description of the relationship between these various non-phonological senses of the word, as well as of the “birth” of words, see Kristó (this volume.)
Turning back to synchronic compounds, there is also some segmental evidence suggesting that the terms of compounds are considered separately by the phonology of English. Compare the underlined ñ’s in anteater, won’t eat it, and antique. Notice that the pronunciation of the underlined ñ in anteater, the last segment of the first term of a compound, is very ‘weak’, it may even be realized as an alveolar flap (that ð- or ð-like sound whose IPA symbol is [r]), and definitely it cannot be heavily aspirated, similarly to the final segment of a word when a vowel-initial word follows it in connected speech (as in won’t eat it), and in contrast to what one expects word-medially, as in antique, where the ñ is followed by a stressed vowel (just as in the previous cases) and consequently is heavily aspirated and cannot be flapped. What this illustrates is again the fact that compounds do not pattern with ‘words proper’, that is, with p-words, but rather the terms are analysed by the phonology as separate p-words.

In Hungarian, one of the phenomena which serve our present purposes is vowel harmony, whose function is, among others, to dictate the frontness of a suffix vowel so that it fits the frontness of the stem to which it is attached. So, a two-form suffix like -val/-vel ‘with’ will take the front form when the (final) stem vowel is front, e.g. lessel ‘with watching/ambush/offside’, egérel ‘with mouse’ and Ágnessel ‘with Ágnes’, but the back form is selected when the (final) stem vowel is back, e.g. fogóval ‘with pincers/pliers/forceps’, madárral ‘with bird’, or when the final front stem vowel is transparent and the preceding vowel is back, e.g. Lacival (and not *Lacivel) ‘with Laci’ and the alternative form Ágnessal ‘with Ágnes’.

What is interesting is the fact that whenever compounds with terms containing vowels of different frontness are formed, e.g. egérúgó ‘mouse trap’ or madárles ‘birdwatching’, the suffix will harmonize with the last term and ignore the rest (egérúgóval, *egérúgóvel; madárlessel, *madárleslessal), and, very importantly, the last term can never be transparent even if it contains a vowel which would qualify as such (compare madárles, with only one possible form, and Ágnes, with two). Clearly, while the phonological structure of Ágnessel coincides with the morphological one, namely, [[Ágnes]sel], this is not the case with madárlessel, where the morpho-syntactic behaviour suggests [([madár][les]]sel), the phonology, however, indicates [madár][lessel].

But not only are there examples of the p-word being smaller than the grammatical word, as in the case of compounds; it can also be illustrated that sometimes it stretches beyond the grammatical word boundary and includes elements which originate in separate syntactic positions. The most obvious such creations are contractions, both ‘grammatical’ (including auxiliary- and *not-contraction) and ‘informal’ (the one most extensively treated in the linguistics literature being wanna-contraction). In the syntactic structure of Peter’s very clever, Peter and is clearly head two separate phrases, however, following the contraction of the two elements, the phonological behaviour of ‘s will be determined by the phonetic makeup of the final segment of its ‘host’, Peter in this example, (i.e., it will assimilate to it in the form of a /z/ here, but a /s/ in Jack’s not so clever), and the resulting construction will consequently be homophonous with the plural of Peter (or Jack). Therefore, Peter’s and Peters exhibit the same phonological behaviour, due to the fact that both comprise single p-words. Such contractions are frequently analysed as cases of cliticisation, that is, as a process.

---

2 The alternation of the first consonant of the suffix, /v/, characterizes a closed set of /v/-initial suffixes of Hungarian, and is beyond the scope of the present discussion, so it will be ignored.

3 There have been cases in Hungarian of former compounds losing that status, similarly to the English examples above; the vowel of the original second term can be transparent in such cases, which is an indication of the fact that the word is no longer treated as a compound, e.g., férfti ‘man’, which gives férftivel if the second vowel is transparent but férftivel if it is not (cf. the alternatives in the case of Ágnes) due to the irregular behaviour of fi.

4 For more information on Hungarian vowel harmony, see, e.g., Kiefer (1994) in Hungarian or Siptár and Tőrkenézy (2000) in English; on the p-word being the domain of vowel harmony in Hungarian, see, among others, Booij (1999) and Nespov and Vogel (1986).
whereby a phonologically dependent object (‘s, ‘ve, n’t, etc. called the clitic) attaches to its host (the element to its left in these cases).\(^5\)

The clitic-like behaviour of certain other function words of English, e.g. the weak forms of prepositions and pronouns, is evident from phonological processes such as "word-initial" flapping and h-dropping. The initial t of the unstressed preposition to is ‘weak’ and frequently flapped in a phrase like lie to me (analogously to a single word like lighter – cf. footnote 8 below), and the initial h of the unstressed personal pronoun him can be deleted in beat him (i.e., beat ‘im), similarly to vehicle or shepherd discussed above. We are led to the conclusion that all of lie to, lighter, beat ‘im, vehicle and shepherd undergo phonological rules in the same way because they constitute single p-words.\(^6\)

In sum, it has been shown above that the phonological sense of the word word is in no way coextensive with any of the other senses; sometimes it covers less, sometimes it covers more.

Finally, I will go on to the discussion of the roles of the p-word in phonology in general, which can be summarised as follows (main points taken from Booij (1999: 48)). Firstly, it serves as the domain of phonotactic constraints: it has been shown above how a consonant cluster like /pb/, well-formed across p-words (e.g., in a hypothetical compound cupboard having the literal meaning ‘a board for cups’, or within a phrase like keep beating) is done away with within the limits of a p-word (in cupboard ‘a closet with shelves where dishes, utensils, or food, and not only cups, are kept’). The more general principle underlying the cupboard-case is the well-known fact that not all combinations of well-formed syllables yield a well-formed p-word, i.e., although both cup and board are existing (that is, licit) monosyllables, they cannot be combined into a single licit p-word without any modifications. Also, the attempt at joining the apparently well-formed right edge /kst/ of a syllable like text with the apparently well-formed left edge /str/ of a syllable like strip will result in the string /kststr/ unattested p-word-internally.\(^7\) Several authors have suggested, on the basis of like phenomena, that the domain of phonotactic constraints is the word, rather than the syllable, as is traditionally thought.

Secondly, the p-word is considered to be the domain of stress assignment rules. We have already seen examples of this, too, since English compounds illustrate that each p-word needs a metrical head, i.e., some degree of stress on one of its syllables, that is why no reduction is found in the second terms of rainbow or lifestyle. Also, there are certain grammatical words which do not usually form p-words on their own but combine with other words (the clitics). Most monosyllabic function words behave in this way, e.g. the preposition (or infinitival marker) to. A remarkable characteristic of these function words is that their unmarked (i.e., neutral) form does not contain a metrical head, i.e., their vowel in unstressed. This is readily explained if they are conceived of as non-p-words, and as such, entities not undergoing the stress assignment rules.\(^8\)

---

\(^5\) Some linguists have even suggested that the clitic(s) and the host in fact form a new type of constituent, the so-called Clitic Group (e.g., Nespor and Vogel (1986)), a view I do not share for reasons whose discussion lies beyond the scope of the present paper.

\(^6\) Therefore, there are no cases of word-initial flapping or h-dropping, since the flapped t’s and the dropped aitches always occupy non-(p-)word-initial positions.

\(^7\) Observations like this have led a number of phonologists to describe the special status of word edges with reference to an obligatorily peripheral constituent (often referred to as the Appendix) which lies outside the scope of basic syllabification.

\(^8\) Independent evidence of the non-p-word status of function words comes from the way they participate in phonological processes, e.g. in flapping, mentioned several times above. Recall that the initial t of unstressed to is ‘weak’ and readily flaps in the relevant dialects in a phrase like lie to me (cf. the single p-word lighter), whereas the beginning of p-words is a ‘strong’ phonological position in that t’s at the left edge of lexical content words are aspirated and never undergo flapping (e.g. the underlined t in tomato both in isolation and in phrases like a tomato, buy tomatoes). The phonological difference between lexical and function words is further elaborated on in Balogné (2002).
Finally, the p-word systematically serves as the domain of phonological rules. We have seen that the domain of application of vowel harmony in Hungarian is the p-word. In English, several segmental alternations have been analysed along similar lines, e.g. \textit{v}-deletion and palatalisation (Hayes 1989, cited in Nespor and Vogel 1986: 150)\footnote{Hayes (1989) takes the Clitic Group as the domain of application of these rules. Cf. footnote 5.}. The rule of \textit{v}-deletion deletes the final /v/ of words if they are followed by any of the clitic-like elements described above (e.g., \textit{leave me alone}; cf. contracted \textit{gimme} from \textit{give me}), but it does not apply if the following word projects a p-word itself (e.g., the underlined \textit{v} in \textit{leave Maureen alone} is not deleted). Occasional cross-word palatalisation is the name of the process which produces palato-alveolar /RIEND/ out of the combination of a word ending in one of the alveolar obstruents /z t d/ and a clitic-like function word beginning with /j/ (e.g., \textit{you}, \textit{your}, \textit{yet}, plus a few other common words including \textit{year} and \textit{usual}): \textit{don’t you see} /dauntju 'si:/, \textit{mind your head} /'mamдзо 'hed/ (for more examples, see, e.g. Shockey 2003: 45). Again, the process is claimed not to take place between separate p-words.

Looking through the arguments for the existence of a separate phonological constituent called the phonological or prosodic word, which takes the morphological word as its starting point but diverges from it in both possible directions (i.e., it may be larger or smaller), we can conclude that the difficulties of defining the notion of the word, so familiar to the native speaker, arise exactly because the different components of grammar cannot be clearly separated but they interact with each other in ways which are hidden in the unconscious knowledge of speakers, and which can only be revealed by systematic scientific study.

References
Kristó, László (this volume) Words and meanings: arbitrariness, listing and lexicalisation.