

The morph as a minimal linguistic form

MARTIN HASPELMATH

MPI-SHH & Leipzig University (December 2019)

Abstract

This paper makes a terminological proposal: that the old term “morph” (which has mostly fallen out of use) can be used for a minimal linguistic form. Many linguists (not only morphologists) need such a term, because we often refer to minimal linguistic forms, but the various terms used by morphologists in roughly this meaning are either unclear or do not refer to forms. The term “morpheme” has three rather different meanings, and other terms such as “vocabulary item” are too abstract. The term “morph” can be used as the basis for defining other widely used terms such as “root”, “prefix”, and “suffix”, which are currently often defined as kinds of “morphemes”. It can also serve as the basis for a clearer definition of suppletion (involving suppletive morph sets) and morph variants, thus avoiding the confusions surrounding the term “allomorph(y)”.

1. The morph: A basic term of general morphosyntax

In this short paper, I make a terminological proposal for general linguistics: The term *morph* should be used for a minimal linguistic form (i.e. a minimal pairing of syntacticosemantic content and a phonological string), because this is a very useful concept that needs a term, and there is no other term that would be suitable. It is the basis for the definitions of *affix*, *prefix*, *suffix*, *root* and other frequently used terms, and is thus a very important basic term.

All linguists are familiar with the term *morpheme*, and this term is sometimes (or even often) used in the sense of a minimal form, but it has two other prominent senses that are fairly different, so it is not suitable as an unambiguous term for a minimal form.

I emphasize that this paper is meant as a methodological contribution, and makes no larger claims, either about how particular morphosyntactic patterns should be described or how they could be explained. I merely make the presupposition that a few basic technical terms should be used consistently, just as a set of basic sounds are rendered consistently by IPA symbols throughout the discipline. Implicitly, this presupposition is widely shared, because many authors use technical terms without defining them, and this would not work unless some basic terms were understood in the same way by everyone. (Consistent terminology is sometimes thought to be more difficult to achieve in morphosyntax than in segmental phonology, but there is no intrinsic reason for this; see Haspelmath 2020.)

Of course, many phenomena in particular languages are special and need special terms (such as “weak declension” in German, or “soft mutation” in

Welsh, or “conjunct order” in Algonquian). Here I am dealing with terms and concepts of general linguistics that could in principle be applied in any language, and that are indispensable for comparative purposes. The definitions of general terms therefore cannot make any reference to language-particular phenomena.

2. Examples of morphs

Before discussing other terms and the research tradition, let us consider a range of concrete examples of morphs, and examples of the use of the basic term *morph* in definitions of other important terms. Examples of morphs are given in (1)-(3) (with glosses in brackets), where some of the morphs are linked to adjacent morphs by hyphens (reflecting the conventional spelling of several adjacent morphs as a single word).

(1) English

<u>cat</u> -s	[cat-PL]
<u>play</u> -ed	[play-PST]

(2) Japanese

<u>neko</u>	[cat]
<u>kan</u> -da	[bite-PST]

(3) Yauyos Quechua (Shimelman 2017: 34-35)

<u>allqu</u> -kuna	[dog-PL]
<u>yatra</u> -rqa	[live-PST]

Examples of technical terms that are defined in terms of ‘morph’ are *root* and *affix*. Two possible definitions are given in (4) and (5). In the examples in (1)-(3), the roots are underlined, and the affixes are not underlined.¹

(4) A root is a morph that denotes a thing, an action, or a property.

(5) An affix is a bound morph that is not a root and that cannot occur on roots of different root-classes.

¹ The morphs in (1)-(3) are all roots or affixes, but languages may also have morphs that are neither roots nor affixes, often called “particles” or “clitics”. Some elements called “determiners”, “adpositions” or “auxiliaries” also fall in this group. There is no widely used general term for non-affix non-root morphs.

My proposed definitions of *root* and *affix* may raise objections, but I will not defend them here (see Haspelmath 2012 for the definition of *root*). I assume that the least controversial aspect of the definitions is that affixes and roots are kinds of minimal forms, i.e. kinds of morphs. This is the point that is important for the current paper: We need the term *morph* in order to have definitions of other terms that we use all the time and that should have a consistent meaning across the discipline.

In the literature, roots and affixes are instead often treated as kinds of “morphemes”, as illustrated by the quotations in (6).

- (6) a. an affix is a “bound morpheme that cannot function as a word on its own” (Booij 2005: 9)
- b. “Bound morphemes are usually classified into roots and affixes. Roots convey lexical meaning and affixes provide additional specification.” (Aikhenvald 2007: 38)
- c. “an affix is a morpheme that is attached to a word stem to form a new word or word form” (Wikipedia in 2019, under “affix”)

It seems that in these definitions, when the authors say “morpheme”, they mean a minimal form, i.e. what I suggest should be called *morph*. But the term “morpheme” has at least three different meanings in the current literature, so it is hardly suitable for the sense of a minimal form that is required to define roots and affixes. I will discuss these meanings in the next section.

It may be noted that the proposal to use the term *morph* in this sense is not quite novel, and indeed, the term is over 70 years old (see n. 4). But it has rarely been used in the intended sense in the last half century, so it makes good sense to remind linguists of a useful term that can serve to give clear and simple definitions of other terms that they use all the time.

3. The term “morpheme”

The term *morpheme* was coined by Jan Baudouin de Courtenay in 1880, and has become widely known through its use in Bloomfield’s *Language* (1933). However, it has not been used consistently over the years. The inconsistent usage was described in some detail by Mugdan (1986) for the first hundred years of the term’s existence,² and it seems to me that the situation has not significantly improved over the last three decades. The term can (at least) be used in one of the three senses in (7), and the senses are often not properly distinguished.

² Matthews (1991: 120) comes to a similar conclusion (apparently independently of Mugdan’s paper, which he does not cite): “The term ‘morpheme’ has a complex history and it is not always clear in which sense it is being used... Present usage is confused.”

- (7) a. morpheme 1:
a minimal form (= a morph)
- b. morpheme 2:
a set of minimal forms with identical syntacticosemantic content
(= a set of homosemous minimal forms, see n. 3)
- c. morpheme 3
a minimal element of (morpho-)syntactic representation

The first sense can be found in definitions of types of morphs, like affix and root (as seen in the preceding section), but it is also widely found elsewhere in the literature. When a linguist needs to refer to a minimal form outside of morphological theorizing, they are very likely to call it *morpheme*. Thus, the first sense could be described as the non-technical (or colloquial) sense of the term.

The second sense is quite prominent in the literature dealing specifically with morphology, and often appears in discussions of “allomorphs”, e.g.

- (8) a. “Allomorphy is the phenomenon that a morpheme may have more than one shape.” (Booij 2005: 31)
- b. “Variant forms of the same morpheme are called allomorphs.” (Kroeger 2005: 289)

For example, Kroeger (2005: 288-289) lists pairs of “allomorphs” such as these:

- | | | | | |
|-----|----|---------|-----------------|---------------|
| (9) | a. | Korean | <i>cekkun-i</i> | ‘enemy-NOM’ |
| | | | <i>oppa-ka</i> | ‘brother-NOM’ |
| | b. | English | <i>liv-ed</i> | ‘live-PPCP’ |
| | | | <i>tak-en</i> | ‘take-PPCP’ |

Each of these examples shows two different affixes that have the same syntacticosemantic content (the same grammatical meaning, or the same morphosyntactic features; we can also say that they are *homosemous*).³ Kroeger treats them as “allomorphs” of the same morpheme, and this usage is not uncommon (see §8 below). The morpheme in this sense is thus an abstract entity: it is not a form, but a set of minimal forms (in other words, a set of morphs) (this set-based definition is clearest in the detailed discussion

³ The term *homosemy* was proposed by Harley (2006: 146) to refer to the full identity of meaning or syntacticosemantic content (in contrast to *synonymy*, which is typically used for semantically similar word pairs such as *big/large*, or *stop/cease*, which are not in complementary distribution).

provided by Mel'čuk 2006: 384-397).

If one uses the term in this second sense, one can no longer felicitously say that a morpheme has a certain shape (e.g. "the morpheme *-i*"), or that some meaning is expressed by a morpheme, or that a word is divided into morphemes, or that several morphemes occur in a certain order, because the morpheme in this sense is not a concrete form. Of course, linguists say these things all the time, but when they do, they necessarily use "morpheme" in sense 1 (i.e. in the sense of 'morph'). But this sense is incompatible with sense 2, and it is not coherent to say at the same time that a morpheme has a concrete shape and that it has several different "allomorphs" (like *-en/-ed* in English, cf. (9b)). Thus, one has to choose between one of the meanings if one wants to use consistent terminology.⁴

Since it is same syntacticosemantic content that is shared by the forms, some authors call this content itself the *morpheme*, as in the following quotation:

"Clearly the past tense form of *loved* consists of two morphemes, the verb-stem *love* and a grammatical morpheme which we can call *Past*, and it's not too hard to draw a line between them. But the past-tense form *took* must likewise consist of two morphemes, the verb-stem *take* and the morpheme *Past*, yet this time we can't draw a neat line at all: the two morphemes are just wrapped up in a single bundle, and we have to appeal to a more abstract level of representation to show that *took* is really *take* plus *Past*." (Trask 1999, key concepts)

Such "abstract morphemes" ("minimal elements of (morpho-)syntactic representation", sense 3 above) have recently become prominent in the Distributed Morphology (DM) tradition, following Halle & Marantz (1993), and Embick's (2015) general book about DM is titled *The morpheme*.⁵ As Mugdan (1986: 36) notes, the notion of abstract morphemes actually goes back to the 1960s (e.g. Bierwisch 1967; Chomsky 1965). In the 1960s and 1970s, they were often alternatively called "formatives", but this term seems to have largely gone out of use.

To summarize, traditions that use the term "morpheme" can treat the two English plural forms *book-s* and *ox-en* in three different ways: They can say (i) that *-s* and *-en* are two different (but homosemous) morphemes expressing plural meaning (sense 1), (ii) that *-s* and *-en* are members of the same English morpheme (sense 2; designated {Plural}, or {-s, -en}, using curly brackets for

⁴ The problem of inconsistency was also present in the early usage of the term "morpheme" by Baudouin de Courtenay and by Bloomfield (as noted by Mugdan 1986). It was solved in the Bloomfieldian tradition through the introduction of the distinction between "morpheme" (for the set of forms, meaning 2) and "morph" (for the minimal form, meaning 1) by Hockett (1947) (an earlier proposal was Harris's (1942) "morpheme alternant"). For some reason, this distinction was not taken up by most of the literature since the 1970s, although it is still found in Matthews (1974; 1991) and Mel'čuk (2006).

⁵ However, Harley (2014: 228) prefers the term *feature bundle* to (abstract) *morpheme*.

sets), or (iii) that there is an abstract morpheme [PLURAL] that may be realized by different exponents (-s and -en) (sense 3)

Because of this multiplicity of meanings of the term “morpheme”, I do not think that the term can be salvaged for future use in technical contexts. It will no doubt continue to be used colloquially (mostly in sense 1), but it seems best to avoid it in technical usage if one wants to be understood more widely. In its colloquial sense, it can be easily replaced by *morph* by authors who value precision and a broad readership.

We began in §2 with definitions of the terms *affix* (which is the basis for *suffix* and *prefix*) and *root*. Having simple definitions of these is important, because the terms *suffix* and *prefix* are very widely used, and in fact my original motivation for writing this paper was that I was very dissatisfied with many of the definitions of these commonly used terms in the literature.

Could one perhaps define an affix as a kind of “morpheme” in sense 2 or 3 in (7)? It seems clear that the answer is no, because we would not want to say that there is a “Korean suffix -i/-ka”, implying that -i is the same suffix as -ka, or that English has “a suffix -en/-ed” (whereas we COULD say that -i and -ka BELONG to the same morpheme {Nominative}, in sense 2, or that they realize the same abstract morpheme [NOM], in sense 3).⁶ Thus, affixes and roots are generally understood as special kinds of morphs, not as “morphemes” (unless one uses this term in sense 1, to refer to morphs).

4. More on forms

The definition of *morph* that was given above contains the term *form* (“a morph is a minimal linguistic form”), which is not remarkable (because linguists talk about forms all the time), but which nevertheless deserves some discussion.

A form is a recurrent pairing of content and segmental shape: “Every language consists of a number of signals, *linguistic forms*. Each linguistic form is a fixed combination of signaling-units, the *phonemes*... We assume that each linguistic form has a constant and definite meaning” (Bloomfield 1933: 158). The elements in (10) are forms of the English language.

- (10) a. *The cats are not on the mat.*
 b. *I’m hungry.* (Bloomfield 1933: 158)
 c. *the cat*
 d. *cats*
 e. -s

Sentence fragments such as *not on the* are not forms because they do not have

⁶ Likewise, using the term “allomorph” (see §8), we would not say that Korean -i and -ka are “allomorphs of the same suffix”.

a coherent meaning, but all sentences, clauses, phrases, words and morphs are also forms (or *expressions*; a *form* and an *expression* are synonymous forms/expressions).⁷

It is unfortunate that the basic term *form* is homonymous with another basic term: *form* in contrasts such as “meaning/content vs. form”. But the two can be distinguished grammatically in English (and other Germanic and Romance languages), because a *form* in the sense ‘expression’ is a count noun, while *form* in the sense of ‘shape’ is (at least optionally) a mass noun. So we can say that a form is a pairing of meaning and segmental form (or, more perspicuously, a pairing of content and segmental shape).

In addition to segmental shape, suprasegmental shape (pitch, intensity, length) is of course important in all languages as well. And syntacticosemantic content may also be expressed by vowel apophony or consonant mutation, by apotony, or by reduplication, as illustrated by the pairs in (11).

- (11) a. English *foot* – *feet*
 b. Welsh *plant* ‘children’ – *blant* ‘children!’ (address form)
 c. Ancient Greek *erootoomen* ‘we ask’ – *eerootoomen* ‘we asked’
 d. Latin *currit* ‘runs’ – *cu~currit* ‘ran’
 e. Beria *húru* ‘saddle’ – *hūru* ‘saddles’⁸

Such nonconcatenative pairs of forms have often been highlighted in order to demonstrate that a simple concatenative view of morphology is insufficient (e.g. Hockett 1947; Anderson 2015: §1.4). However, such cases do not have any consequence for the definition of the terms *a form* and *a morph*. Vowel and consonant changes, suprasegmental changes and reduplication are normally thought of as PROCESSES, not as FORMS, and hence not as morphs. Moreover, morphologists sometimes say that a “morpheme” may be realized by zero, but again, it should be clear that a form cannot be zero (because it is a pairing of a shape and a meaning). Thus, morphs (as forms) must be segmental.

The term *morphology* may suggest that the term *morph* should be used in a wider sense, including more than segmental forms. However, nobody has suggested that morphology should be defined as the study of morphs. Morphology was originally defined as the study of linguistic forms, but as syntax became more prominent, morphology became restricted to “word forms” (though these are difficult to delimit from phrasal forms; Haspelmath 2011). It seems that nobody would deny that concatenating morphs is what

⁷ To be sure, term *expression* is much more often used for longer forms/expressions (perhaps comprising several words), while the term *form* tends to be used in particular for word-forms (e.g. 10d). But I am not aware of any definitions of these terms that restrict them in these ways.

⁸ Jakobi & Crass (2004: 108); Beria is a Saharan language.

many morphological and syntactic rules do (or can be thought of as doing),⁹ so even if the term *morphology* is not quite transparent, is it not completely inappropriate either.

In addition to nonconcatenative aspects of shape, linguists have also sometimes identified meaningless strings, sometimes called “empty morphs” (e.g. *-th-* in English *far-th-er*). Such strings are not morphs by the current definition unless they are said to have some (morpho)syntactic content (e.g. providing a suitable stem for an affix to attach to).¹⁰ But note that the definition given here does not require that all elements of an utterance belong to some morph (just as not every morph is either an affix or a root, cf. note 1). Analysis must eventually be complete, but technical terms need not cover everything.

Thus, morphs are segmental, as they are kinds of forms, which are defined as having a segmental shape. Nonsegmental aspects of shape do of course also have an important role in morphosyntax, but they will have to be captured by other terms (not discussed in this short paper).¹¹

5. Minimality

A morph is defined as a minimal form, which means that it does not consist of other morphs.

Now of course, whether a form is minimal or can be further divided into smaller forms with their own content is not always clear. Linguists often disagree about morphological segmentation, e.g. whether English *him* and *them* can be analyzed as containing an accusative suffix (*hi-m*, *the-m*) or should be treated as unanalyzable forms, like *her* and *us*. Such questions will never go

⁹ Some linguists, such as Anderson (1992; 2015), prefer to look at all of morphology in terms of rules rather than minimal forms. For purposes of describing particular languages, one may get this programme to work, but for general and comparative morphology, we need terms such as *root* and *suffix*, and these must be defined as kinds of forms. As noted, it is often unclear where phrases end and words begin, and even Anderson would say that the concatenation of forms has a role to play in phrases.

¹⁰ One colleague has suggested to me that I am not doing justice to the literature on “pure morphology”, which emphasizes that morphs are pairings of segmental shape and morphological content. But linguists do not have a consistent general way of distinguishing between morphology and syntax (cf. Haspelmath 2011), so the term “semanticsyntactic content” is intended as a cover term for semantic and morphosyntactic content.

¹¹ Mel'čuk (2006) defines morphs as “minimal signs” (see also Mugdan 2015: §2.2), but he allows a variety of different phenomena to count as “signs” (including vowel change phenomena), so that his definition ends up being highly complex, and his “morphs” are hard to delimit. (It is not clear, for example, whether word order alternations such as question-word fronting should not also be regarded as “elementary linguistic signs”, whereas it is very clear that such operations/transformations are not forms.) In general, it seems that the only practical way to achieve terminological unification is to have straightforward definitions for simple terms. If one wants to propose more complex concepts, one should choose novel or more complex terms for them.

away, and they will always be answered differently by different linguists (and presumably also differently by different speakers of a language).¹² But the existence of such borderline cases does not affect the definition of the term *morph*.¹³

What is important to realize is that minimality is not the same as “simplicity”. Linguistic forms are sometimes said to be “complex” even when they cannot be segmented into two or more morphs. Famously, Bloomfield (1933: 161) defined the *morpheme* as a “simple form”:

“A linguistic form which bears no partial phonetic-semantic resemblance to any other form, is a *simple* form or *morpheme*.”

But morphs such as English *feet* or Welsh *blant* (cf. (11a-b) above) are not “simple” in this sense, because they are semantically and formally similar to other forms. In English, there are at least a few other pairs like *goose* – *geese* and *woman* – *women*, and in Welsh the pattern is even productive. Thus, these forms are generally regarded as “morphologically complex”, but they cannot be segmented,¹⁴ so they are distinct minimal forms (i.e. distinct morphs).

Of course, the contributions made by the suprasegmental and segmental changes seen in (11) are in an important way parallel to the contributions made by grammatical markers (= grammatical morphs), and one eventually needs to have a way of talking about this parallel (e.g. in terms of morphological constructions, or realization rules). But recall from §1 that I am not trying to make a contribution to methods of elegant description, let alone theories of mental morphosyntactic knowledge. This paper is about clear terminology, and in this context, a notion of a “process morph(eme)” would not be compatible with terminological clarity. A “process” is not a form, and since a morph is a form, it cannot be a process.

6. Morph variants

¹² Or differently at different speaker ages (e.g. German *Handtuch* ‘towel’ is typically acquired early as a minimal form, but later in life – after learning the orthography – many speakers come to segment it into *Hand* ‘hand’ and *Tuch* ‘cloth’, even though the meaning is not transparent).

¹³ How one segments forms also depends on what one allows as syntacticosemantic content. Segmenting Spanish verb forms into three morphs (stems, stem vowels and endings: *habl-a-mos* ‘we speak’, *com-e-mos* ‘we eat’, etc.) may be justified by saying that the stem vowels are syntactically required. But in cases like English *-ceive* and *-duce* (cf. *perceive/receive/deceive* vs. *perception/reception/deception*, *reduce/deduce* vs. *reduction/deduction*), there is no such justification, so these would not be morphs. (A description of English could or should still segment these, of course, but they cannot be regarded as morphs for comparative purposes.)

¹⁴ Here I leave aside the possibility of segmenting *foot* and *feet* into a root *f-V-t* and a singular and a plural infix, respectively. (Positing a nominative *p-* and a vocative *b-* in Welsh would be even more far-fetched.)

In addition to being clear about forms and minimality, we need to say more about phonological variability. It is well-known that morphs often have phonological variants, as is illustrated in (12).

- | | | | | | | |
|--------|---------|------------|--------|-------------------|--------------|----------|
| (12)a. | German | infinitive | [n] | <i>beiß-en</i> | 'to bite' | [baɪsn] |
| | | | [m] | <i>schreib-en</i> | 'to write' | [ʃʁaɪbm] |
| b. | English | plural | [s] | <i>cat-s</i> | | |
| | | | [z] | <i>dog-s</i> | | |
| | | | [ɪz] | <i>bitch-es</i> | | |
| c. | Turkish | plural | [ler] | <i>ev-ler</i> | 'houses' | |
| | | | [lar] | <i>kitap-lar</i> | 'books' | |
| d. | Polish | 'leg' | [nog] | <i>nog-a</i> | 'leg' | |
| | | | [nodz] | <i>na nodz-e</i> | 'on the leg' | |
| e. | Spanish | 'want' | [ker] | <i>quer-ér</i> | 'to want' | |
| | | | [kjer] | <i>quíér-o</i> | 'I want' | |

In these cases, it is perfectly normal to say that [s] and [z] in English are the same suffix, and that [ker] and [kjer] in Spanish are the same root. Thus, linguists do not think that an affix or a root (and hence a morph) must always have exactly the same shape. However, the shape variation must be of a phonological nature. If two different segment sequences are not phonologically similar at all, they will be regarded as different morphs even if they are homosemous (semantically equivalent), as we saw in (9a-b). This applies both to affixes and to roots (see §7 below on suppletive affixes and roots).

As in the case of minimality (§5), linguists will often take different decisions concerning the distinction between phonological variants of the same morph and different morphs. When an alternation is very regular and occurs in a variety of different environments, it will normally be treated as phonological, even if it is not very natural phonetically (this applies to the Polish alternation between [g] and [dz], for example). When an alternation applies only to few cases but is phonetically natural, it is also usually treated as phonological (as in the case of English [s]/[z]/[ɪz]). When it is neither phonetically natural nor regular (e.g. English *is* vs. *are*), then everyone treats the two segment sequences as different morphs, even if the forms are descended from the same morph at an earlier stage. Thus, there is very broad agreement on the principle of distinguishing between morph variants and different homosemous morphs (see §7 below), even if in many individual cases, judgements may differ.

In §3 above, I said that English *-en* and *-ed* (in 9b) would not be said to be the

same suffix. By contrast, I think that we WOULD say that *-s* and *-es* (in *dogs* and *bitches*) are the same suffix, or variants of the same suffix, just as we would say that English *analyze* is the same word as *analyse*, and that both are spelling variants of the same verb. This identity relationship might be underlined by deriving both *-s* and *-es* from the same abstract underlying form, but this is not necessary in order to recognize their identity (after all, there is no abstract underlying form for the spelling variants *analyse/analyze* either).

7. Suppletive morph sets

A morph is sometimes homosemous (see note 3) with another morph and occurs in a complementary morphosyntactic distribution. For example, Latin has the root *fer-* ‘carry’ in the Present tense, but *tul-* ‘carry’ in the Perfect:

(13) Latin

- a. *fer-o* ‘I carry’, *fer-s* ‘you carry’, etc.
- b. *tul-i* ‘I carried’, *tul-isti* ‘you carried’, etc.

Such sets of morphs are called suppletive morph sets. Linguists usually say that such cases involve different roots, so *fer-* and *tul-* are different morphs.¹⁵

Suppletive morph sets that occur in complementary distribution are also found in inflection. For example, German has a variety of different plural morphs, which are sometimes conditioned by gender or derivational suffix, but are often distributed arbitrarily, e.g. *Tag/Tag-e* ‘day(s)’, *Lied/Lied-er* ‘song(s)’, *Staat/Staat-en* ‘state(s)’, *Wrack/Wrack-s* ‘shipwreck’. In Hungarian, the second-person singular form of verbs usually contains the morph *-sz*, but the homosemous morph *-el* when the verb stem ends in a sibilant (Carstairs 1988):¹⁶

(14) Hungarian

1SG <i>kér-ek</i>	‘ask’	<i>keres-ek</i>	‘look for’
2SG <i>kér-sz</i>		<i>keres-el</i>	
3SG <i>kér</i>		<i>keres</i>	

Again, in such cases linguists usually say that there are different affixes (we already saw examples of this phenomenon in (9a-b) above). Thus, suppletion

¹⁵ E.g. Booij (2005: 34): “The Italian verb *andare* has two suppletive roots: *and-* when the root is not stressed, and *vad-* when the root is stressed in the verbal paradigm (*vádo*, *andiámo*, etc.).”

¹⁶ Cases like (14) show that it is important not to confuse *phonological variants* with “phonologically conditioned variants”. Two phonological variants of the same morph are phonologically similar, ideally in a very regular way, but phonological conditioning may apply to suppletive morph sets (such as *-sz/-el*) as well (cf. Paster 2009).

involves different morphs, i.e. a set of different (but homosemous) forms.¹⁷

The term *morpheme* would not be appropriate for the different elements in sets like *fer-/tul-* and *-sz/-el*, because it has often been used in an abstract sense (sense 3 in 4c), to refer to whatever content the suppletive morphs share (e.g. Lyons 1968: 182-183: “*worse* is composed of two morphemes, one of which it shares with *bad*, and the other of which it shares with *taller*, *bigger*, *nicer* etc.).

As noted in the preceding section, the boundaries between suppletion and (morpho)phonological alternations are not sharp (cf. Kiparsky 1996). In *bad/worse*, we clearly have a suppletive morph set, and in cases like *lo[n]/lo[ŋ]er*, we clearly have two variants of the same morph. But is *bett-* (in *better*) the same morph as *be-* (in *best*)? Or are they two different morphs, constituting a suppletive morph set? As in the case of minimality in §5, different linguists (and different speakers) will always give different answers,¹⁸ but what matters here is that there is a clear conceptual distinction: Either two shapes are variants of the same form, or they are two different forms.¹⁹

8. “Allomorphy”: Supplemorphism and variomorphism

The term “allomorph(y)” is used in scare quotes in this paper, because I do not think that it can be defined in such a way that the definition is very largely compatible with existing usage. Like the term “morpheme” (§3), the term “allomorph” has the problem is that it is used in rather different ways in the literature (though here I will highlight two rather than three different senses).

In the morphology textbooks, “allomorph” tends to be used in the sense of ‘morph variant’ (§6). But in the research-oriented literature, it tends to be used in the rather different sense of ‘member of a suppletive morph set’ (and the term “allomorphy” is then used in the sense of “suppletion”).

For example, Lieber (2009: 158) says in her textbook that “allomorphs are phonologically distinct variants of the same morpheme”, and Booij (2005: 31) defines “allomorphy” as “the phenomenon that a morpheme may have more

¹⁷ One might question the suppletive nature of a morph set on the grounds that the distribution is not completely complementary (e.g. Swedish *bättre* ‘better’, which is the usual comparative counterpart of *god* ‘good’, but it exists alongside a form *godare* with a somewhat different meaning). But again, this does not affect the definition of *suppletive*, because it is due to some vagueness in the more general term *complementary distribution*, which is an independent fact.

¹⁸ Occasionally one hears the form *bestest*, which suggests that *best* is not analyzed into *be-* and *-st* (i.e. that *good/bett-/best* are a suppletive morph triple).

¹⁹ In the literature, the notion of “suppletive allomorphy” is strongly associated with inflectional markers and roots. By contrast, for synonymous derivational markers (e.g. *warmth* vs. *roughness*; authors typically talk about “suffix rivalry” (e.g. Arndt-Lappe 2014). It seems that the stereotypical expectation is that homosemous inflectional markers may not coexist side by side but must occur in complementary distribution, but it is unclear whether there is such a general distinction between inflection and derivation (cf. the existence of “overabundance” in inflection; Thornton 2011).

than one ... particular phonological form" (likewise Harley 2006: 131).²⁰ The main issue that these textbooks mention is that some linguists limit the term "allomorph" to unpredictable morph variants. Thus, Booij (2005: 32-33) says that one might want to limit the term to cases like (12d-e) (e.g. Spanish *quer-/quier-*), where the alternation is morphophonological (not phonologically automatic), presumably because the morphological notion of "allomorph" is not really needed when there is clearly a unique underlying form from which the phonological realizations can be derived by phonological rules.²¹

In the research-oriented literature, by contrast, the term "allomorph" is typically used for a member of a suppletive morph set. Kiparsky (1996) contrasts "allomorphy" with "morpho-phonology" (cf. also Carstairs 1987), and in the same vein, Bonet & Harbour's (2012) survey paper discusses suppletive morph sets in contrast to (morpho)phonological morph variants, noting that "others use the term 'suppletion' to refer to multiple underlying forms" (2012: 199).²²

Thus, two rather different meanings are associated with the same term "allomorph": 'morph variant' and 'suppletive morph'. In view of this situation, it seems best to avoid the term (just as it is best to avoid the term "morpheme") and to use other terms instead.²³

Maybe one could use *supplemorph* for a situation where there is a suppletive morph set, and *variomorph* for a situation where a morph has several variant shapes.²⁴ But while one could talk about several homosemous *supplemorphs* (e.g. English *go* and *wen-*, two homosemous morphs meaning 'go'), the term **variomorph* makes no sense (the morph variants *-[d]*, *[t]* and *[ɪd]* of the English plural morph *-s* are not different morphs).

²⁰ As expected given its prevalence in the textbooks, this meaning has also found its way into Wikipedia (2019): "An **allomorph** is a variant form of a morpheme, that is, when a unit of meaning varies in sound without changing the meaning."

²¹ From the perspective of this paper, this difference is irrelevant, because all that matters is whether two phonological shapes can be regarded as variants of the same form (just as *analyse* and *analyze* are spelling variants of the same verb). Clearly, the relationships between morph variants vary in regularity, but this plays no role for the definition of this term.

²² The terminological confusion is particularly striking in Harley & Noyer's (2003) survey article, where they first say that they want to "distinguish allomorphy from suppletion" (p. 463), but then that "there are two different types of allomorphy: suppletive and morphophonological" (p. 474).

²³ Another reason why "allomorph" is not a good term is that it is quite opaque. Originally, the terms *morpheme/morph/allomorph* were meant to be analogous to *phoneme/phone/allophone*, but while an allophone of a phoneme is a phone, an allomorph is almost never said to be a kind of morph.

²⁴ Note also that the members of a set of *supplemorphs* may themselves have morph variants, e.g. in the English morph set *-ed/-en* (two different homosemous past participial morphs), the first morph has the variants *[d]*, *[t]* and *[ɪd]*. Such situations cannot be described well by an "allomorph" concept.

9. Homonyms and idioms

Above I defined a morph as a minimal pairing of meaning and form, or content and shape. I already discussed two ways in which such minimal forms are not straightforward to delimit from other forms: Their minimality may be in question (e.g. in the case of English *them*, possibly segmentable into *the-m*), and their identity may be in question (e.g. in the case of *be-* in *best*, which may be identical to *bett-* in *better* or not).

In addition, of course, one may wonder whether a pair like *bank* and *bank* is a single polysemous morph or whether we are dealing with two different morphs that just happen to have the same shape (i.e. two homonyms). The same issue arises with grammatical markers: Is the Russian *-a* in *volk-a* ‘wolf (accusative)’ the same morph as *-a* in *volk-a* ‘wolf (genitive)’ (so that we are dealing with systematic syncretism), or are they two different homonymous morphs? These questions have often been discussed, but like the questions about minimality and identity, they will always be answered differently. Again, fortunately, they are not relevant for the definition of ‘morph’.

Less often discussed is the question of how to deal with complex but noncompositional expressions like *white wine*, *hard disk*, *deadline*, *spill the beans* ‘divulge a secret’, *leave no stone unturned* ‘search intensively’, or German *Handtuch* ‘towel’, mentioned in note 11). Strictly speaking, these cannot be said to consist of a combination of morphs, but in many cases, their figurative nature is quite transparent to the speakers (who may not even notice the lack of compositionality, e.g. in the case of *white wine*). In those cases where there is no transparency at all (e.g. *kick the bucket* ‘die’, *hot dog* ‘sausage sandwich’), it is harder to justify a treatment in terms of multiple morphs, but such cases seem to be uncommon. As in the other cases, I have nothing to say on the issue of delimitation here, which arises in any event and does not seem to be resolvable in a general way.

10. Related terms: *Exponent*, *vocabulary item*, *listeme*

Finally, let me briefly discuss three terms that are similar in meaning to *morph*: *exponent*, *vocabulary item*, and *listeme*.

The term *exponent* seems to be fairly old, but has become better known to linguists through Matthews (1972; 1974). In the 21st century, its popularity seems to have increased, and there are now books with “exponence” in their title (Trommer (ed.) 2012; Harris 2017). A typical use of the term *exponent* is as a minimal form that realizes a set of morphological features. However, given the context of inflectional paradigms, where some cells often lack forms that express the relevant inflectional meaning, many authors work with a notion of a “zero exponent” (e.g. Plank 1999: 282; Calabrese 2011; among many others).

Moreover, in the discussion of “multiple exponence”, both a plural suffix and plural vowel change are said to be exponents (e.g. German *Hals* ‘neck’, *Häls-e* ‘necks’, Harris 2017: 1). According to the definition proposed here, a morph (or any other form) cannot be nonsegmental, so a vowel change is an “exponent” which is not a morph.

The term *vocabulary item* has also become widely used, in the Distributed-Morphology literature, and it seems to have basically the same meaning as *exponent*. While Harley & Noyer (2003: 468) say that “Vocabulary Items provide the set of phonological signals available in a language for the expression of abstract morphemes”, they include in their examples a vocabulary item that has no phonological features, and is thus not a form (or a morph) in the sense of the present proposal.

Finally, the term *listeme* (originally proposed by Di Sciullo & Williams 1987) is used extensively in Harley’s (2006) textbook in a sense similar to ‘morph’. Harley takes pains to explain that English plural *-s/-es* represent the same listeme, while plural *-i* (as in *alumn-us/alumn-i*) is a different listeme. But if a listeme is ‘anything that we need to memorize’ (and that thus needs to be listed in some mental repository), then this term is much broader than ‘morph’. Thus, idiomatic compounds like *newspaper* and *wallpaper* need to be memorized (Harley 2006: 100), and are thus listemes, but they are at the same time very transparent, so one would not say that they are morphs (as discussed in §9). As Jackendoff (1997: Chapters 5-6) discusses extensively in the context of multi-word expressions, the relationship between what is listed mentally and what is described by rules is not straightforward, even if we assume (counterfactually) that our linguistic knowledge is stored nonredundantly. Thus, a notion of listeme is certainly different from *morph* as defined here.²⁵

11. Concluding remarks

This paper has discussed a range of different terms that are used widely in morphosyntax, using a text genre that is almost unknown otherwise (but see Mel’čuk 1982; 2006; Mugdan 1986; 2015 for similar work on technical terms). But in view of the fact that the terminology that we use is often inconsistent and unclear, I hope that my readers find this exercise useful.

Many morphologists prefer to direct their energies into rather different directions, e.g. attempting to describe particular phenomena in a highly elegant way (and hoping that the approach will eventually generalize), or proposing ambitious architectures intended to reflect the mental organization of linguistic knowledge in all speakers and all languages. If we made good

²⁵ Harley (2006: 85) characterizes listemes as “minimal meaningful units of sound”, but later she posits a zero listeme (for the plural of *sheep*; 2006: 148), so that her use of listeme seems to be no different from current uses of the terms *exponent* and *vocabulary item*.

progress in these efforts, we might dismiss terminological discussion as a minor and irrelevant aspect, because true discoveries should be apparent even without clear terms. But if progress in the field is not expected to be fast, then we might as well devote some energy to consistent terminology, so that at least the definitions of basic terms in our textbooks (and thus on Wikipedia) will be non-circular and clear. Without clear terminology, we will keep talking past each other, and if we make real progress, we will not be able to communicate this clearly.

Finally, I should remind the readers that I am not making claims about how languages are best described. I defined ‘morph’ in terms of a segmental string, because everyone agrees that such entities may be useful for comparing languages, not because I would suggest that morphs are sufficient to describe languages (they are obviously not sufficient). Neither do I want to suggest that all minimal strings that are usefully segmented are morphs (see the brief discussion of English *-ceive* in n. 12). While general theories must ultimately be complete, this methodological paper makes no attempt to provide a complete set of concepts (see also the brief discussion in n. 1 of roots, affixes, and other elements for which there exists no general term). It merely provides a simple and clear definition of the term *morph* (as a minimal linguistic form), for the purposes of general linguistics, which will probably be useful for many linguists.

References

- Aikhenvald, Alexandra. 2007. Typological distinctions in word-formation. In Timothy Shopen (ed.), *Language typology and syntactic description*, Vol. 3, 1–65. Second Edition. Cambridge: Cambridge University Press.
- Anderson, Stephen R. 1992. *A-Morphous morphology*. Cambridge: Cambridge University Press.
- Anderson, Stephen R. 2015. The morpheme: Its nature and use. *The Oxford handbook of inflection*, 11–33. Oxford: Oxford University Press.
- Arndt-Lappe, Sabine. 2014. Analogy in suffix rivalry: The case of English -ity and -ness. *English Language & Linguistics* 18(3). 497–548. doi:10.1017/S136067431400015X.
- Bierwisch, Manfred. 1967. Syntactic features in morphology: General problems of so-called pronominal inflection in German. *To honour Roman Jakobson*, 239–270. The Hague: Mouton.
- Bloomfield, Leonard. 1933. *Language*. New York: H. Holt and Company.
- Bonet, Eulàlia & Daniel Harbour. 2012. Contextual allomorphy. In Jochen Trommer (ed.), *The morphology and phonology of exponence*, 195–235. Oxford: Oxford University Press.
- Booij, Geert E. 2005. *The grammar of words: An introduction to linguistic morphology*. Oxford: Oxford University Press.
- Calabrese, Andrea. 2011. Investigations on markedness, syncretism and zero exponence in morphology. *Morphology* 21(2). 283–325. doi:10.1007/s11525-010-9169-y.
- Carstairs, Andrew. 1987. *Allomorphy in inflexion*. London: Croom Helm.

- Carstairs, Andrew. 1988. Some implications of phonologically conditioned suppletion. In Geert Booij & Jaap van Marle (eds.), *Yearbook of morphology 1988*, 67–94. Dordrecht: Foris.
- Chomsky, Noam A. 1965. *Aspects of the theory of syntax*. Cambridge, MA: MIT Press.
- Embick, David. 2015. *The morpheme: A theoretical introduction*. Berlin: De Gruyter.
- Halle, Morris & Alec Marantz. 1993. Distributed morphology and the pieces of inflection. In Kenneth L. Hale & Samuel J. Keyser (eds.), *The view from Building 20*. Cambridge, MA: MIT Press.
- Harley, Heidi. 2006. *English words: A linguistic introduction*. Malden, MA: Blackwell.
- Harley, Heidi & Rolf Noyer. 2003. Distributed Morphology. *The second Glot International state-of-the-article book*, 463–496. Berlin: De Gruyter Mouton. doi:10.1515/9783110890952.463.
- Harris, Alice Carmichael. 2017. *Multiple exponence*. New York: Oxford University Press.
- Harris, Zellig S. 1942. Morpheme alternants in linguistic analysis. *Language* 18(3). 169–180.
- Haspelmath, Martin. 2011. The indeterminacy of word segmentation and the nature of morphology and syntax. *Folia Linguistica* 45(1). 31–80.
- Haspelmath, Martin. 2012. How to compare major word-classes across the world's languages. In Thomas Graf, Denis Paperno, Anna Szabolcsi & Jos Tellings (eds.), *Theories of everything: in honor of Edward Keenan* (UCLA Working Papers in Linguistics 17), 109–130. Los Angeles: UCLA.
- Haspelmath, Martin. 2020. Towards standardization of morphosyntactic terminology for general linguistics. vol. (to appear in an edited volume).
- Hockett, Charles F. 1947. Problems of morphemic analysis. *Language* 23(4). 321–343. doi:10.2307/410295.
- Jackendoff, Ray S. 1997. *The architecture of the language faculty* (Linguistic Inquiry Monographs 28). Cambridge, MA: MIT Press.
- Jakobi, Angelika, Joachim Crass & Bakhit Seby Abdoulaye. 2004. *Grammaire du beria (langue saharienne): Avec un glossaire français-beria*. Köln: Köppe.
- Kiparsky, Paul. 1996. Allomorphy or morphophonology? In Rajendra Singh & Richard Desrochers (eds.), *Trubetzkoy's orphan*, 13–31. Amsterdam: Benjamins.
- Kroeger, Paul. 2005. *Analyzing grammar: An introduction*. Cambridge: Cambridge University Press.
- Lieber, Rochelle. 2009. *Introducing morphology*. Cambridge: Cambridge University Press.
- Lyons, John. 1968. *Introduction to theoretical linguistics*. Cambridge: Cambridge University Press.
- Matthews, Peter H. 1972. *Inflectional morphology*. Cambridge: Cambridge University Press.
- Matthews, Peter H. 1974. *Morphology*. Cambridge: Cambridge University Press.
- Matthews, Peter H. 1991. *Morphology*. 2nd edition. Cambridge: Cambridge University Press.
- Mel'čuk, Igor' Aleksandrovič. 1982. *Towards a language of linguistics: A system of formal notions for theoretical morphology*. München: Fink.
- Mel'čuk, Igor A. 2006. *Aspects of the theory of morphology*. Berlin: De Gruyter. <https://doi.org/10.1515/9783110199864>.
- Mugdan, Joachim. 1986. Was ist eigentlich ein Morphem? *Zeitschrift für Phonetik, Sprachwissenschaft und Kommunikationsforschung* 39(1–4). 29–43. doi:10.1524/stuf.1986.39.14.29.
- Mugdan, Joachim. 2015. Units of word-formation. In Peter O. Müller, Ingeborg Ohnheiser, Susan Olsen & Franz Rainer (eds.), *Word-formation: An international handbook of the languages of Europe*, vol. 1, 235–301. Berlin: De Gruyter Mouton. doi:10.1515/9783110246254-017.
- Paster, Mary. 2009. Explaining phonological conditions on affixation: Evidence from suppletive allomorphy and affix ordering. *Word Structure* 2(1). 18–37. doi:10.3366/E1750124509000282.

- Plank, Frans. 1999. Split morphology: How agglutination and flexion mix. *Linguistic Typology* 3. 279–340.
- Shimelman, Aviva. 2017. *A grammar of Yauyos Quechua* (Studies in Diversity Linguistics 9). Berlin: Language Science Press. <http://langsci-press.org/catalog/book/83>
- Thornton, Anna Maria. 2011. Overabundance (multiple forms realizing the same cell): A non-canonical phenomenon in Italian verb morphology. In Martin Maiden, John Charles Smith, Maria Goldbach & Marc-Olivier Hinzelin (eds.), *Morphological autonomy: Perspectives from Romance inflectional morphology*, 358–381. Oxford: Oxford University Press.
- Trask, R. L. 1999. *Key concepts in language and linguistics*. London: Psychology Press.
- Trommer, Jochen (ed.). 2012. *The morphology and phonology of exponence*. Oxford: Oxford University Press.