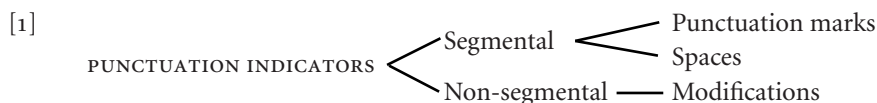


1 Preliminaries

1.1 The domain of punctuation

The central concern of punctuation is with the use of the various **punctuation marks**, such as the full stop, comma, semicolon, colon, question mark, quotation marks, parentheses, and so on. These serve to give indications of the grammatical structure and/or meaning of stretches of written text. The punctuation marks are all **segmental** units of writing – i.e. they fully occupy a position in the linear sequence of written symbols. There are, however, various **non-segmental** features which can serve the same kind of purpose as the punctuation marks. For example, titles of literary or other works may be italicised as an alternative to being enclosed in quotation marks. And while the end of a sentence is indicated segmentally by a punctuation mark (a full stop, question mark, or exclamation mark), the beginning of a sentence is indicated non-segmentally by capitalisation of the first letter. We will therefore regard punctuation as covering the use not only of punctuation marks but also of such non-segmental features as italics, capital letters, bold face, and small capitals. Ordinary lower-case roman represents the default form, and these non-segmental features can be regarded as **modifications** of the default form.

One other important aspect of punctuation is the use of space, notably to separate one word from the next. Space between words is a segmental unit: like the punctuation marks, it occupies the whole of one position in linear sequence. For example, in this sentence a word space occupies the fourth position, the thirteenth position, and so on. We will use the term **punctuation indicator** as a general term covering punctuation marks and the other devices that fall within the domain of punctuation. The classification is thus as follows:



On another dimension, we need to clarify the domain of punctuation with respect to the size of the unit to which the punctuation applies. The punctuation marks mentioned above generally occur within a sentence (including its final boundary) but outside the individual words. There are two punctuation marks, however, that are normally word-internal: the apostrophe and the hyphen. Words may also contain various non-segmental marks, **diacritics**, but we do not regard these as falling within the domain of punctuation. For example, accents (which do not of course appear in native English words, but are nevertheless found in some words that are otherwise fully anglicised, such as *fiancé*)