What are Determiners?

The category “determiner” is a grouping broader than the traditional notion of the article, and encompasses not only the articles (*a, and, the*), but also other items that can fill the article slot. This generality applies to linguistic approaches to determiners, including ours: "the set of determiners has been defined as a set of closed-system items that are mutually exclusive with the articles." (Quirk et al. 1972:139). Linguistic analyses of determiners necessarily delve into the count vs mass distinction in nouns (*this pen* vs *some ink*), and how singular vs plural distinctions affect which types of nouns allow which types of determiners (*many pens* is correct/acceptable; *many ink* is not).

Our coding for determiners focuses on number agreement and mass/count restrictions, to characterize the syntactic behavior of determiners. *Many* and *most* have “plur” in their variants slot; *either* and *each* have “sing,” for example. Determiners occurring only with mass (uncount) nouns will have the variant “uncount,” e.g. *less* and *much*. Some determiners can occur with either singular or uncount nouns (e.g. *this* and *that*), and are coded “singuncount,” while others occur with either plural or uncount nouns (e.g. *more* and *other*), and are corde “pluruncount.” Those determiners occurring without number restrictions (e.g. *the*, *which*, *some*) are coded “free” in the variants slot.

There are many differences in subcategories and details among the many linguistic treatments of determiners. For example, Quirk et al group Predeterminers (*all, both, half; double, twice, three times* (etc); *one-third, one-fifth* (etc)) Ordinals and closed-system Quantifiers and call them "closed system premodifiers." These kinds of differences occur throughout linguistics and likely other academic fields. Quirk et al have chosen to divide into many subcategories, what we have grouped into the single determiner category (setting aside such minor details as the lack of Lexicon records for *one-third* and similar). They are mutually exclusive, and have distinctive singular/plural agreement behavior that adjectives lack.

The following are the 38 determiners with UMLS Lexicon records as of July 2017: *a, all, an, another, any, both, certain, each, either, enough, every, few, fewer, last, less, many, more, most, much, neither, no, other, several, some, such, that, the, these, this, those, thy, what, whatever, which, whichever, nary a, nary an, suchlike.*

Determiners vs Adjectives:

Perhaps because dictionaries are written to be used by the average educated person, they follow the practice of traditional grammars in their various POS assignments of the words in our list of determiners. *A*, *an* and *the* are usually called articles; the rest are usually termed adjectives. Therefore, we will address the differences between determiners and adjectives, and why we see the “determiner” as a valid POS.

First, let’s consider what adjectives are. Quirk et al (Chapter 5) follow tradition in even considering adjectives (and adverbs) to be a POS.

Four general characteristics of adjectives:

1. Can freely occur in attributive position, e.g. *the **happy** children*
2. Can freely occur in predicative position, e.g. *This book is **old***
3. Can be premodified by the intensifier *very*, e.g. *The children are very **happy***

4. Have comparative and superlative, either by inflection (*-er/-ier –est/-iest*) or periphrasitically (using *more* & *most*).

That said, not all words traditionally considered adjectives will have all four characteristics; our Lexbuild adjective records accommodate these four characteristics while not making them obligatory. Here, for example, are the Lexbuild records for *happy, generous,* and *stearic*:

```
{base=happy
entry=E0030812
        cat=adj
        variants=reg
        variants=inv
        position=attrib(1)
        position=attrib(3)
        position=pred
        compl=infcomp:subjc
        compl=fincomp(o)
        nominalization=happiness|noun|E0030811
}


{base=generous
entry=E0029555
        cat=adj
        variants=inv;periph
        position=attrib(1)
        position=pred
        nominalization=generosity|noun|E0029554
        nominalization=generousness|noun|E0587596
}


{base=stearic
entry=E0206172
        cat=adj
        variants=inv
        position=attrib(3)
        stative
}
```

You may note that our record for *happy* includes more information about its syntactic behavior—that as a pre-noun modifier, it may occur either before (position=attrib(1)) or after (position=attrib(3)) a color adjective (color adjectives are position 2), that it can take particular types of infinitival or finite complements, and that it has the nominalization *happiness*. The adjective *generous* does not have as much syntactic information in its record, as it does not have commonly associated complements. Neither *happy* nor *generous* is stative, because these adjectives do not refer to a characteristic state of the noun they modify, but rather are under the conscious control of the (usually) person of whom they are said. The test for stativity is that if you can say, “Don’t be so [adjective]!” then that adjective is not stative. The adjective *stearic* is a stative adjective, as it would be very odd to admonish someone not to be so stearic. The majority of biomedical adjectives are stative. You may note that the record for *stearic* has less information than the other examples given. It only occurs immediately pre-noun, never as a predicate adjective, has no characteristic complements, and is stative. The most common array of

syntactic information in Lexbuild adjective records, has slightly more than *stearic*, as most adjectives can occur as predicate adjectives, e.g.:

```
{base=behavioral
spelling_variant=behavioural
entry=E0012256
        cat=adj
        variants=inv
        position=attrib(3)
        position=pred
        stative
}


{base=arthroscopic
entry=E0408054
        cat=adj
        variants=inv
        position=attrib(3)
        position=pred
        stative
}


{base=presurgical
spelling_variant=pre-surgical
entry=E0319382
        cat=adj
        variants=inv
        position=attrib(3)
        position=pred
        stative
}
```

Our approach to adjective position is overall simpler than many other linguistic analyses in its focus on whether or not the adjective can occur before or after a color adjective. However, unlike many other linguistic anayses of adjectives, we also mark adjectives for their potential to occur as predicate adjectives. We believe this approach to the syntactic behavior of adjectives addresses well the pragmatic considerations of NLP applications. Moreover, not considering determiners to be adjectives avoids unnecessary complication in their description, while better describing their syntactic agreement behavior.

Determiners vs Possessive Pronouns:

The traditional group of possessive pronouns (*my, your, her, his, our, their*) occupy a sort of middle ground between pronouns and determiners. Quirk et al call them determiners; we call them pronouns. As pronouns, they are marked as possessive, and as to their gender. Our treatment of them as pronouns is not only to follow their traditional label as "possessive pronouns," but also to more easily describe (code) the 1st, 2nd & 3rd person variants, which are parallel to the (other) personal pronouns. Syntactically, they are mutually exclusive with determiners, so in that respect, Quirk et al are on solid

ground in calling them determiners. Because they have 1st, 2nd and 3rd person variants, we chose to treat them as pronouns.

Conclusion:

Our extended list of determiners aims to enumerate those items that can fill in the "determiner" position in a mutually exclusive way. Further, the agreement restrictions shown by determiners and not adjectives, is linguistically defensible, and may be pertinent to certain NLP analyses.

Reference

Quirk, Randolph, Sidney Greenbaum, Geoffrey Leech and Jan Svartvik. (1972) *A Grammar of Contemporary English*. London: Longman.