SupMat_2: Annotation examples of HK and SR

In this SupMat, we provide some examples of the more intricate ways in which HK and SR combine, leading to various types of reference. While in the majority of cases the values for HK and SR align, both being present or absent, different combinations are possible. We will first discuss the often-generic combination **HK+ SR-**, followed by **HK- SR+** and finally the non-referential **HK- SR-**.

The combination of **HK+ SR-** typically occurs when we are dealing with either a generic use of a term, which is activated by concept, but can also occur in case of a non-generic use, which is activated by the preceding discourse. We provide an example of genericity in (1):

(1) <ø> Capitalism will brook no delay.

In non-generic **HK+ SR-** situations, establishing whether a referent (regardless of whether the noun refers) is known to the hearer or not depends on other elements in the sentence or wider context around the target noun. In our dataset we found a number of such instances, notably in a document that appears to be a flying manual: we provide such an example in (2):

(2) Since the majority of ø modern machines have ø airbrakes which are held closed by a geometric lock, it is essential that the method of checking them must be foolproof for this kind of ø system.
First open the brakes or ø spoilers fully, checking them visually whenever this is possible.
Then close them with a firm push on the lever in ø order to establish the habit.
The pilot should be able to feel the geometric lock operating just after <**the**> airbrakes are flush with the wing surface.

In this example, the NP *the airbrakes* has an SR-, but HK+ reading. The referent is known to the hearer as it has been previously mentioned in the text (first line), but because these are guidelines that apply to *the majority of modern machines*, the referent cannot be considered specific in a strict sense; it therefore appears to stand somewhere between specific and generic. Although there is no specific referent, the referent does not stand for the whole category either and is thus not fully generic.

A concept noun like *capitalism* is very likely to be HK+ SR- in most contexts. This is also true of nouns referring to sports for example as in *She plays football/tennis/golf.* However, there are nouns that can be used both as generic and non-generic, and this is entirely context dependent. Let us illustrate this point with an example that is **HK- SR+**:

(3) She then put <ø> **parsley** on top of the mashed potatoes in my plate.

In this example, the noun is not used to refer to the concept of parsley but to an actual portion of it. The portion of parsley is, in this context, unknown to the hearer, but specific (HK- SR+). However, the noun *parsley* can also be used in a generic sense: *I love parsley,* where reference is made to the concept rather than to a portion. Nouns such as *golf, tennis,* and *capitalism* are particular in that they usually only ever refer to concepts. Note that even when used in a narrower context, e.g., *I am currently playing golf*, the noun is used to refer to the concept of golf and not to a portion of it (HK+ SR-).

Another example of a referent that is **HK- but SR+** can be found in (4) below. In this case, the speaker/writer makes reference to a particular act, an 'adoption act' but does not expect the hearer/reader to be aware of this particular act, since it is labelled as 'new'. The specific reference is

clear, as there is a description of what the act changed and a specific date for when it came into force but there is no evidence that the reader/hearer is expected to know about this act.

(4)   It was 1974 before that particular form of discrimination was removed by <a> new adoption act.

Finally, the example in (5) illustrates a referent that is **HK- and SR**-. In this example, no generic statement is made about the concept 'picture' either. Rather the speaker uses a non-specific member of the category 'picture' as a referent: there is no reference to a specific picture, but instead to some unspecified picture that stands for the category of 'pictures that are not illustrated'. The typical example of this distinction is given by the difference between *She wants to marry a millionaire – but he doesn't seem too keen* and *She wants to marry a millionaire – any millionaire will do.* In the first case, the NP refers to a specific millionaire, whereas in the second it does not: all she wants is to marry a member of the category 'millionaire'.

(5)   This is as unsatisfactory as reading about <a> picture which is not illustrated – although an instance is activated it's not a specific, identifiable picture