The "givenness hierarchy"
Looking primarily at points made by Gundel et al. (1993).

1 Cognitive statuses for referring expressions

(1) The Givenness Hierarchy
   a. In focus; *it*
   b. activated; *that*, *this*, *this N*
   c. familiar; *that N*
   d. uniquely identifiable; *the N*
   e. referential; indefinite *this N*
   f. type identifiable; *a N*

"Each status on the hierarchy is a necessary and sufficient condition for the appropriate use of a different form or forms."

This is supposed to be an **implicational hierarchy**, meaning that if a referring expression has the status, e.g., "familiar", it also—necessarily—has the statuses "uniquely identifiable," "referential," and "type identifiable."

1.1 Type identifiable

The addressee is able to access a representation of the type of object described. Goes well with indefinite *a*.

(2) I couldn't sleep last night. **A dog** (next door) kept me awake.

"Thus, *a dog* in (2) is appropriate only if the addressee can be assumed to know the meaning of the word dog and can therefore understand what type of thing the phrase *a dog* describes." Since "type identifiable" is the least restrictive option, this presumably means that a phrase like *a schmerinkterpan* is never appropriate. Could be true, I suppose.

1.2 Referential

The speaker intends to refer to a particular object or objects. The addressee needs to be able to access or construct the speaker's intended referent. Goes well with indefinite *this*.

(3) I couldn't sleep last night. **This dog** (next door) kept me awake.

The distinction between "referential" and "type identifiable" seems to be extraordinarily subtle. But so is the distinction between *this dog next door kept me awake* and *a dog next door kept me awake*. (What do you think? Does it seem like you can get away without constructing a referent in (2)? In (3)? They say that studies have shown that *this dog* is more likely to be followed up by further discussion of the dog than *a dog* is. . .)

1.3 Uniquely identifiable

The addressee can identify the intended referent on the basis of the noun alone (without the rest of the sentence it appears in). Goes well with *the*.

(4) I couldn't sleep last night. **The dog** (next door) kept me awake.

1.4 Familiar

The addressee is able to uniquely identify the intended referent because he's already got a representation in (possibly long-term) memory. Goes well with demonstrative *that*.

(5) I couldn't sleep last night. **That dog** (next door) kept me awake.

(*This dog* is different, that requires not just a familiar referent, but also an activated one. Something to do with its meaning of "proximal," one would imagine.)

1.5 Activated

The addressee is able to uniquely identify the intended referent because he's currently got a representation in short-term memory. Goes well with noun-less demonstrative *that*.

(6) I couldn't sleep last night. **That** (next door) kept me awake.

Cf. also *This dog (we're discussing) kept me awake.*

1.6 In focus

The referent is not only in short-term memory but is the center of attention. Goes well with noun-less demonstrative *it*.

(7) **It's** the same dog that bit Mary last summer.

I'm not a fan of the name of this status. In fact, this may be a source of some confusion. This is not the way we will use the word "focus" for much of the rest of the course. This is much more like what we will call "topic" in the rest of the course.

2 Inferrables

The idea here is that sometimes things can be evoked by being related in terms of some kind of world knowledge. E.g., "you've left out a whole paragraph." Or "John entered a room. The window was open."

(8) I dropped ten marbles and found all of them, except for one. It's probably under the sofa.

(9) # I dropped ten marbles and found only nine of them. It's probably under the sofa.

Much of this section of the paper is partly just connecting with some prior literature on the subject, and trying to make the case that the categories in the present paper are enough to handle "inferrables" as well.

3 Universality of the Givenness Hierarchy

For this, I'll turn to the paper itself rather than reproduce bits here. Table 1 shows forms for Chinese, English, Japanese, Russian, and Spanish. All of the categories discussed have an English form associated with them. All of the languages other than English have fewer status–form correspondences. All distinguish "in focus" from "activated"—but that's about all that is common across all five languages. From the other direction, there is some commonality of function for demonstratives, definite articles—they land in pretty much the same category, at least, across languages.

4 The Givenness Hierarchy and natural language discourse

Before this section, there wasn't really much energy spent motivating the claimed *implicational* nature of the hierarchy. But the idea is supposed to be that a referent that has the status of "activated" it necessarily also has the status of "familiar," etc. So, at the most extreme, something that is in focus should be *unconstrained* as to what form is used for it.

But it's decidedly weird to say "Here, see this pencil? Let's talk about this pencil." followed by "John likes a pencil" (to mean something like "John likes it."). So, what evidence *is* there for this hierarchy?

Gundel et al. (1993) look at "naturally occurring discourse," classify referents as to their cognitive status based on the evolving conversation, and count the forms.

Looking at their tables, it does seem that essentially, the lines spread down from the most restrictive licensed form. But they are "top heavy"—most of the time, the most restricted form that can be used is used.

5 The Givenness Hierarchy and Grice's Maxim of Quantity

(10) a. Simplified English disallows the use of passive, progressive, and perfective auxiliary verbs, among other things.
b. It requires engineers to break up long compound nouns and technical expressions into chunks of three or less [sic] elements.
c. This requires engineers to break up long compound nouns and technical expressions into chunks of three or less [sic] elements.

(11) Maxim of Quantity
a. Make your contribution as informative as required (for the current purposes of the exchange).
b. Do not make your contribution more informative than required.

(12) I ate some of the cookies.

(13) I have two televisions.

In this context, the idea goes something like *you'd said "a dog" so you must be unable to uniquely identify it, because if you could have uniquely identified it, you would have said "the dog" (or one of the other options)*. Because "the dog" is available in fewer situations, you'd prefer to use it if it is applicable.

So, for (10c), the use of *this* suggests that *it* wasn't appropriate (since you would have used *it* if you could have). *It* would have allowed continuing the topic (as in (10b)), so *this* must indicate a new topic in (10c).

As for why you don't *always* seem to use the one closest to *it* on the scale, they suggest that (particularly for definites), the distinction doesn't make much difference to the meaning so the choice is freer than among indefinites.

For next time. *Note that this is a change in the schedule*. Look at the Ghomeshi et al. (2004) paper, read up to but not including section 4. That is, up to about two-thirds of the way down p. 336.

References

Gundel, Jeanette K., Nancy Hedberg & Ron Zacharski. 1993. Cognitive status and the form of referring expressions in discourse. *Language* 69(2): 274–307.