Approved for release by NSA on 12-01-2011, Transparency Case# 63852

You've Got to Go by the Rules

JACK GURIN

An interesting and sometimes amusing discussion of the use of the definite article in our native language.

Most native speakers of English are unaware of the rules that control the language they speak. If they have ever thought about the question, they probably feel that while there probably are rules, they are likely to be outnumbered by exceptions. How else can you explain, for example, that we say "lead a child by *the* hand," but "an article made by hand?" Or "go to school" but "go to *the* store?" Merely examining the question may induce a sensation approaching panic, as in the case of the centipede who, when asked how he knew which leg to put down when, promptly fell into the ditch in confusion.

The truth is that our speech is really governed by strict rules, although here and there some flexibility is allowed. In spite of our picture of ourselves as ruggedly individualistic and proud of our opportunities to make our own choices, we unquestioningly submit to these rules of speech. Suppose you were entertaining a neighbor who, as far as you knew, was a native-born American. During a conversation in which he talks about his children, he says, "My oldest is really laid back about school. He takes it all in stride, gets good grades, and still manages to play *the* baseball with the team." Without giving it a second thought, you know that English is not this man's native language, no matter how accent-free his speech nor how fluent his use of slang. He gave himself away by violating a rule that no American over the age of five, no matter how poorly educated, would ever fail to observe: ONE NEVER USES THE DEFINITE ARTICLE WITH THE NAME OF A GAME. Can you imagine asking "Anyone for *the* tennis?" Or admitting that you enjoy an occasional game of *the* poker?

What about people who violate the rules, even though their native language is English? What about those who speak "incorrectly"? Some large segment of the population says "between you and I," or "times ain't what they was." The phrase "I couldn't care less" has been transformed into "I could care less" by many, retaining the same meaning when it really should signify the opposite. What has happened here is that these people are following rules, too, but they are different rules, judged to be incorrect with regard to what is taken as the standard for acceptable and "correct" English. There are many examples of these substandard practices becoming acceptable and eventually replacing what used to be considered standard.

While coaching some immigrants in the niceties of the English language recently, I was confronted with the question "What are the rules for the use of the definite article in English?" To my dismay I found that I had no answer, and promised to prepare myself for our next session. I consulted several handbooks of English grammar, texts I had left over from high school or college, and discovered that they dismissed the subject after differentiating between the definite article *the* and the indefinite article *a* or *an*. It was clear that the reader was expected to KNOW how to use the articles. I was reminded of a

question raised sometime ago about the definition of "human beings" versus "animals." Of course everyone KNOWS what a human being is, although I am not aware that there exists any useful legal definition, or even a philosophical or theological one. Suppose a missing link were discovered, and its life taken. Would that be murder?

Why worry about whether there are rules or not? We native speakers don't seem to need them, and the immigrants will eventually mimic their teachers and their neighbors enough to be able to communicate more or less effectively. Their children, of course, will be fluent. Anyway, if we confront the adult foreign student of English with a long list of rules, he would probably not be able to absorb them and would continue to make the same mistakes as before.

But what about the machines, those that we have been told will eventually be "intelligent" enough to compose reasonably acceptable prose (or poetry, for that matter). There is no hope that they will be able to originate anything at all understandable unless we set them up with the rules of the language—the ones we all follow without being aware of them, the ones we absorb just by living in this country, hearing, speaking, or reading its language. If we don't provide the rules, the poor stupid machine will simply regurgitate the words and phrases we have supplied to it, in whatever order and form those words happen to be. The result will not only fail to resemble standard English, it may not communicate the desired meaning at all. The machine cannot be expected to KNOW the rules; it must be told, and told with a precision and at a level of detail not given to us in our schools.

To return to the question that started this whole discussion, what are the rules for the definite article in English? Do such rules exist and can they be cataloged? Is it possible to reduce this language of ours to a set of rules, no matter how extensive, that will take care of all eventualities? Logic indicates that the rules have to be there, or else there would be no consistency at all to our language. There have been attempts to list the rules, with varying degrees of completeness. The *American Heritage Dictionary*, for example, presents a list of eight different uses of *the*, not nearly enough to explain the examples that easily come to mind.

The *Oxford English Dictionary* does a much more thorough job of providing meanings for the word *the*. But by far the most impressive listing of rules, together with some explanation of the reasoning behind them, is presented by the Danish philologist Otto Jesperson, author of *A Modern English Grammar* (in seven volumes) and, in a smaller compass, *Essentials of English Grammar* (one volume). Jesperson, who died in 1943, is regarded by many as the most distinguished scholar of the English language who ever lived, with an unmatched love and respect for the subject of his scholarship.

Jesperson explained the nature of the definite article:

> *The* is generally called the definite article: a better name would be the defining or determining article. It has really two distinct functions, that of determining in itself, and that of determining in connexion with a following word or words containing the essential specification. We therefore speak of *the article of complete*, and *the article of incomplete determination*.

Complete determination—when there is no question in the mind of the speaker (and presumably the listener) as to what person or thing is being referred to.

> Shut the door, please.
> Where is the telephone?
> I'm going back to the office.
> And all the while he knew the answer.

Incomplete determination—when the definite article is not in itself sufficient to determine what we are speaking about. Under this category may be reckoned all combinations in

which the substantive has an adjective, as in "the gray horse"; thus very often with superlatives and ordinals, as in "the eldest boy, the fifth girl in the third row," etc. The clearest instances are, however, found when the supplementary determination follows after the substantive, this with a prepositional group and with a relative clause.

The man in the moon.
The book you lent me.
The memorandum I wrote yesterday.

Jesperson goes on to explain that while use of the indefinite article denotes unfamiliarity (or ignorance) with the specific object, it could be a definite member of a class. In the case of the definite article, however, *the* "plus a substantive in the singular denotes one individual (supposed to be) more or less familiar to the speaker or writer: Some image or notion of the thing or person denoted by the substantive is (supposed to be) already found in the consciousness of the speaker or writer before he makes the statement."

Now we can go ahead and list some of the rules for using *the*, taking ideas from both Jesperson and the Oxford Dictionary to help classify those rules. We will also begin by admitting that the list will be incomplete, debatable, overlapping, and arbitrary.

I. Implicit Contextual Basis

After something has been mentioned, we use the definite article for things that are obviously connected with it. For example, after mentioning a university, we may go on to speak of the professors, the courses, the colleges, the library, etc. As soon as we have checked into a hotel, we want to see the room and to find out where the coffee shop (or bar) is located. In a new neighborhood, we want to know who the neighbors are and how good the schools are.

II. Use of *the* to Denote Typical or Preeminent

He certainly looked the officer and gentleman.
Quite the ladies' man, I'd say.

Also, with *something of*, *a touch of*:

He retains something of the professor in his appearance.
She still has a touch of the snob.

Also, with adjectives like *typical* or *perfect*:

The typical Marine DI.
She made the perfect companion.

With stress on the *the*, to denote the typical or perfect specimen:

That's what I'd call *the* dream house.
According to him, World War II was *the* war.

Before the name of a unique object or one so to be considered, or one of which there is only one at a time:

The sun; the earth; the sea; the sky; the Messiah.

With names of natural phenomena, seasons, etc.:

The spring; the summer; the day; the wind; the cold; the north; the east.

III. Some Special Cases

With the names of languages, but only in consciously elliptical phrases, as "from the German" (*language* or *original* understood).

Before names and titles of men, sometimes a corruption of the French *de*:

Robert the Bruce; Sir Simon the Montfort.

With any part of the body of a person previously named or indicated, instead of the corresponding possessive pronoun:

He took him by the hand (i.e., his hand).

Before *body, mind, soul,* or parts, functions, and attributes of these, when used figuratively, and before parts of personal attire:

What appeals to the tongue may not please the stomach.
Where the heart and the mind do not agree.
It took the heart out of him.
He was taken by the seat of the pants and thrown out.

IV. Generic Use of *the* in the Singular

Before the name of a person, animal, or plant, used generically:

The dog is the friend of man.
You can see the woman in the little girl.

With anything used as the type of its class, as with the names of musical instruments, tools, etc.:

Break a condemned man on the wheel.
The cinema is no substitute for the theatre.
He tried but couldn't teach him the violin.
Tolstoy took the novel to new heights.

Symbolic use:

The lower deck doesn't want to give orders; it looks to the bridge for them.
The pen is mightier than the sword.

Names of dances used generically:

He could dance neither the polka nor the mazurka.
She was extremely graceful in the tango.

Names of some general ideas when used generically:

Whether to sit in the shade or the sun.
Women finally won the vote.

Names of some diseases:

Catch the flu.
Avoid him like the plague.

But not others:

He came down with malaria.

V. Distributive Use

With units of measurement:

He bought them by the dozen.
I can watch this by the hour.

VI. Familiar Use

I'll have to check with the missus.
That's not a nice thing to say in front of the kids.

But, in some cases:

Better see Mama every night or you can't see Mama at all.

VII. In Exclamations

It is essential that there be an emotional element in the statement.

The fool! The stupid fool!
The nerve of that woman!

But, with words meaning "nonsense":

Bunk! Baloney! Bureaucratic hogwash!

VIII. Before Comparatives

The smaller the unit the better its performance.
That makes it all the funnier.

IX. Before Superlatives and Ordinals

Superlatives:

I'll take a dozen, if they're the cheapest.

In *of* phrases:

He's in the best of health.

In set phrases:

One hopes for the best.
No one here was in the least to blame.

Ordinals:

On the first of May.

But sometimes you have a choice:

On May the first—or—on May first.

With *most*:

They are great guys, for the most part.
The most you can expect is a thank you.

With *last*. We differentiate between two kinds of usage:

He comes to see us the last Sunday of every month.
He came to see us last Sunday.

With *next*. Again, we differentiate between usages:

The people in the next house.
The people next door.

X. Adjectives Denoting a Class

With an adjective used absolutely, usually denoting an abstract notion:

Distinguish between the known and the unknown.
From the sublime to the ridiculous.

With *all*:

It makes all the difference.

But, with expressions of time:

I slept all night.

Before an adjective or participle having a plural application (usually of persons) :

Alms for the poor.
Rock music has little appeal to the middle-aged.

With the past participle:

No use preaching to the converted.

XI. Proper Names

With a substantive in the plural, chiefly the name of a nation, class, or group of people, where *the* = "those who are," "the. . . .taken as a whole." Also with family surnames:

The Romans left us quite a legacy.
The Joneses are of Welsh origin.

When one substantive is defined by another:

The poet Virgil.

Usually the proper name precedes, especially when the whole phrase becomes a recognized appellation:

William the Conqueror; Harry the Horse.

The is used when a personal name is particularized by an adjective. The adjective usually precedes, but sometimes follows the name, but in either case *the* stands first:

The incomparable Duse; the charismatic Jack Kennedy.

But when the adjective becomes a permanent epithet, *the* and the adjective usually follow:

Alfred the Great; Ethelred the Unready.

River names take *the*:

Floating down the Mississippi.
Exploring the Amazon.

Oceans, straits, channels take *the* :

The Adriatic; the Atlantic.
The Bosporus; the Dardanelles.

But lakes do not:

On the shores of Lake Michigan.
The beauties of Lake Como.

Countries usually do not take *the* unless plural:

The Balkans; the United States; the Indies.

Exceptions to the preceding may be explained by the influence of a foreign language, ellipsis, or perhaps because they were originally river-names:

The Caucasus; the Palatinate; the Saar; the Klondike.

Towns—very few take *the*:

The Hague (from den Haag).

Parks don't need *the*:

Rock Creek Park; Central Park.

Streets, roads, and railroad stations don't take *the*:

Main Street; Benning Road; Grand Central Station.

Ships almost always need *the*:

Sailing on the QE II.
Aboard the Mary B.

XII. More Examples Where *the* Isn't Needed

Direct address:

Luck, be a lady tonight.
Nurse!

Legal terms:

Counsel was doing his best.

Meals:

By that time, dinner was ready.

Institutions:

He saw her again after church.
I'll have to report this to headquarters.

Town:

Go to town; live in town; man about town.

Names of periods and dates:

Sunday; Christmas; July.
Time for school; lunch period.
Winter had gone, and it was spring again.
Nothing happened until morning.

While it would be presumptuous to claim that the preceding provides anything more than a good beginning for the exhaustive listing required to define all the uses of the definite article, it does show the possibilities. Judging from the number of issues surrounding the use of just this one word, the task of compiling the rules for the entire language would appear to be a very large one indeed, even though some of the work has already been done. But if the job appears impossible because of its magnitude, I offer the following. Suppose we had no dictionaries of our language and discovered that we need to list all of its words with their definitions. What a daunting task that would be—if we had to start from scratch. But fortunately that would not be necessary, because for a couple of hundred years—a surprisingly short time, come to think of it—lexicographers have been compiling lists of words and examining examples of their usage. No one expects the compilers of new lexical works to start the process from the beginning. They invariably perch on the shoulders of their predecessors and begin their flights from there. Those who will be called upon to do the job on all the *rules* for the language will not have quite the same advantage, but the job can be done nonetheless. We just have to remember that the poor, dumb machine must be told *everything* about the language, since it does not possess that marvelous innate ability that is the gift provided to every human being, the ability to communicate through whatever language happens to be used in his neighborhood.

We may postpone undertaking this admittedly large and expensive task, whether for English or for any other language for which the need exists, but not for long. Perhaps we will find a way to enlist the aid of the computer itself in performing the necessary analysis. Perhaps we will find compelling reasons for postponing the effort. But there is really no choice in whether we do the job or not; it is just a matter of time.

~~(FOUO)~~ Mr. Gurin, for many years a fixture at NSA until he retired, remains within the Agency but now as an employee of Systems Research Laboratories (SRL), working in R83. He admits to being very excited about the possible roles technology may play in assisting linguistic and analytic personnel in their tasks. The focus of his efforts, he insists, remains the human in the workplace, and the role of technology is to help out by taking over work that is drearily repetitive or too voluminous to be attractive or cost-effective for humans.