

Where Vocabulary Meets Grammar: Verb Subcategorization Errors In ESL Writers¹

- Learners of English as a second language at intermediate and advanced stages have often mastered the majority of the major syntactic constructions in English. Yet, many grammatical errors persist in their writing. A high percentage of these errors, though labeled *grammatical*, do not in fact represent problems with pure syntax but rather mistakes in using given lexical items in constructions they do not belong in. By utilizing concepts from modern transformational-generative theories, the authors trace such errors to incorrect or incomplete lexical subcategorization. The nature of these errors is discussed from both a theoretical and pragmatic perspective, and the major classes of subcategorization errors for English verbs are identified. The article argues that ESL teachers, particularly writing teachers of students beyond the beginner's level, need to be aware of the source of these errors so that they can distinguish them from other types of grammatical errors and more effectively help their students to overcome them.

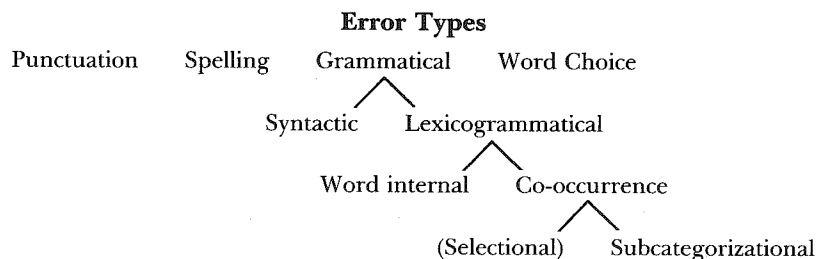
Lexical subcategorization is a significant aspect of modern linguistic theory which has received surprisingly little attention in discussions of ESL student errors. In fact, errors resulting from incorrect or incomplete subcategorization are generally just labeled *grammatical*, even though writing teachers are usually able to recognize them as being somehow different from many other types of grammatical errors, such as those involving word order, number agreement, lack of an article with a countable singular noun, and so forth. We feel that there is some value in writing teachers being conscious of the nature of these errors in lexical subcategorization and the important ways in which they differ from others labeled as grammatical. We begin by discussing these differences and then give a description of 14 major types of lexical subcategorization errors for English verbs with an example of each type. Verbs form the nucleus of sentences and it is with verbs that the widest variety of subcategorization errors occur. Later sections suggest ways of recognizing and dealing with such errors in intermediate and advanced ESL writing. We focus on this group of learners because we believe that although the type of error analysis we propose is applicable in many cases to lower levels

and to oral production at all levels, it is at intermediate and advanced levels of writing that any teacher correction of this type of error is most likely to be successful.

There are a number of different sentence-level error types that ESL writing teachers are confronted with at all levels. These are typically categorized as errors in spelling, punctuation, word choice, and grammar. The last of these categories, grammatical errors, is rather vaguely and inconsistently defined in the field; errors that don't fit into the other categories are often assigned to it.

There are some major distinctions within this general category of grammatical errors, however, that can significantly affect the way the teacher reacts to the errors and leads the student toward their correction. At the most basic level, there is the distinction between purely syntactic errors involving rules of word order and inflection that are governed by the grammatical categories of words or phrases and errors that involve idiosyncrasies of particular words or classes of words within a category. We refer to this second group as *lexicogrammatical* errors because they crucially involve some sort of lexical misinformation that leads to an ungrammatical sentence (for our purposes *ungrammatical* is operationally defined as being recognized by native speakers to have been produced by a nonnative speaker due to some factor other than word choice). Within the category of lexicogrammatical errors itself there are similarly two major divisions: errors internal to a word and errors in co-occurrence with other words. Finally, within the category of co-occurrence errors there are errors that are predominantly semantic in nature (labeled as *violations of selectional restrictions* in the standard theory of transformational grammar [Chomsky, 1965] and those which are seen as more syntactic in nature (labeled as *subcategorization errors* in the standard theory). Because of the semantic nature of the former we assume, as most syntacticians currently seem to, that they are better classified as purely lexical errors or errors in word choice.

Figure 1 shows the error types discussed so far and their relationship in a classificational hierarchy.



The following examples of errors in the hierarchy should help make the differences clearer.

Syntactic errors

- a. I saw a cat black yesterday.
(order of adjective and noun reversed)
- b. The clock that I bought it last week is already broken.
(extra pronoun in relative clause)

Lexicogrammatical errors

- a. Word internal
I wrote a good composition.
(irregular verb)
- b. Co-occurrence
 1. Sectional
A dog roared at me as I entered the yard.
(word choice: *roar* for *bark*)
 2. Subcategorizational
I enjoy to run in the rain.
(wrong form of complement)

In the earlier stages of language learning, many of the errors in all of the above categories are viewed as either developmental errors or errors occurring from the students' attempt to overreach their linguistic competence in the target language. Consequently, many teachers at this level are justifiably selective about which errors they point out to the student, particularly in speaking situations. At more advanced levels, though, and especially in writing classes, there is often a concern on the part of the teacher that errors which were overlooked in earlier stages may become a more or less permanent feature of the student's English unless they are systematically dealt with. At this point, recognizing the nature of their error is vital. Since writing is in many cases an activity that can be monitored, or consciously attended to (in the sense of Krashen's [1982] monitor theory of language learning versus acquisition), the teacher's classification of the error potentially becomes quite significant in leading the student to overcoming it. Specifically, if an error is predominantly syntactic in nature it can be dealt with by referring in some way to appropriately formulated syntactic rules, such as those for adjective-plus-noun word order and the structure of relative clauses. If the error is lexicogrammatical, on the other hand, then idiosyncratic information about individual words must be brought to the student's attention.

Linguistic Theory and Subcategorization.

In the earliest versions of transformational grammar (Chomsky, 1957), grammaticality was based solely on whether or not a sentence could be generated by some combination of phrase structure and transformational rules. Both of these rule systems referred exclusively

to expansions or manipulations of syntactic categories (e.g., noun, verb, verb phrase), allowing the generation of such obviously unacceptable sentences as *Fred put the napkin* and *The napkin was put by Fred*. By the time the aspects model, which was to become the standard theory, was introduced (Chomsky, 1965), the need for referencing in lexical insertion not only the category of the word itself but also the categories of its surrounding constituents had been recognized. This subcategorization frame, as it was called, became a part of the lexical entry for words. In the case of verbs, it allowed the specifications of what categories could follow them. For example, the frame [NP] was specified for transitive verbs while intransitive verbs were given the frame [] with no following object NP.

The past 20 years have seen some dramatic changes in the theories that have evolved from transformational grammar, changes which many ESL teachers trained in the standard theory are unaware of. Of greatest relevance here is the fact that the transformational component has been steadily weakened while the power of the lexicon has been expanded to deal in a principled manner with the distinction between truly general rules and lexical idiosyncrasies. In the process, most transformations have fallen by the wayside: Some current versions of Chomsky's revised extended standard theory (also known as government and binding theory) hypothesize only a single transformation—*move alpha*, or *move anything anywhere*—constrained of course by a number of universal and language specific conditions (Chomsky, 1982; see Radford, 1981 for a less formal introduction to the theory). As a result, the output of many of the classic transformations, such as passive and indirect object movement is now generated more or less directly by the phrase structure rules, and those formerly transformational relationships, such as active-passive, are now accounted for primarily in the lexicon.

Common Types of Lexical Subcategorization Errors With Verbs.

Because of their central importance, the discussion of lexical subcategorization errors below focuses on verbs. It is important to realize, however, that other major categories (noun, adjective, adverb) can be subcategorized as well.

We assume for the moment that whatever the ultimate cause of the error (interference, over generalization, etc.), the immediate cause is the lack of an appropriately specified subcategorization frame for the verb. The list that follows includes most of the common lexical subcategorization errors for verbs. In each case, we first give a correct sentence utilizing the grammatical construction to make it clear that the error is lexical in nature. Asterisks mark the ungrammatical sentences.

1. *Transitives used as intransitives*. Some verbs which semantically imply an object can occur with the object unspecified; others cannot.

- a. *I ate the sandwich in the kitchen./I ate in the kitchen.*
- b. *I put the sandwich in the kitchen./*I put in the kitchen.*

2. *Intransitives used as transitives*. Some intransitive verbs have transitive counterparts, often with the meaning *cause to x*, where *x* is the intransitive verb; others do not.

- a. *The water boiled./The woman boiled the water* (the woman caused the water to boil).
- b. *The baby cried./*The woman cried the baby* (the woman caused the baby to cry).

3. *Incorrect passive of transitive verb*. Nonlinking verbs which are followed by noun phrases in the active voice are usually classified as transitive and can appear in the passive construction; some verbs, especially certain stative ones, cannot.

- a. *John earned \$100./\$100 was earned by John.*
- b. *This book cost \$100./*\$100 was cost by this book.*

4. *Stative verbs in the progressive*. While verbs can generally be used in progressive constructions, most semantically stative verbs are ungrammatical in such constructions.

- a. *I am learning French.*
- b. **I am knowing French.*

5. *Fixed particle moved*. The particle in verb-particle combinations, most of which are to some degree idiomatic, can usually be moved after the object. However, certain verb-particle combinations do not allow this option.

- a. *Fred ran off the cat./Fred ran the cat off.*
- b. *Fred ran into an old friend./*Fred ran an old friend into.*

6. *Dative movement errors*. Dative, or indirect object, movement, where the indirect object moves to direct object position and the preposition *to* or *for* is dropped, is allowed with many verbs that take indirect objects but not all of them.

- a. *Tom told the story to everyone./Tom told everyone the story.*
- b. *Tom explained the problem to everyone./*Tom explained everyone the problem.*

7. *Intrusive be*. The passive auxiliary *be* is sometimes inserted into nonpassive sentences. Though treated as a general grammar error in Burt and Kiparsky (1972) and Richards (1973), it was argued in Hubbard (1983) that while such errors may be essentially syntactic with lower level students, the great majority of these errors occur with a restricted set of intransitive verbs in advanced writers, specifically those

verbs taking nonagentive subjects. In many languages, these intransitives pattern differently from intransitives with agentive subjects and may even take the same morphology as the passive construction. Even in English, there are a few verbs that appear in a passive-like construction which cannot take an agentive *by*-phrase in certain contexts.

- a. *The theater is located/situated (*by Fred) on Elm Street.*
- b. **The strange event was occurred last May./*The problem is exist today in many countries.*

8. *Gerund complement for infinitive.* One of the most common subcategorization errors involves using the wrong complement form after a verb. (Items [9] through [11] also involve complement errors, and there are other types besides these.) In this particular case, a gerund complement is used where the verb requires an infinitive.

- a. *I enjoy seeing you.*
- b. **I expect seeing you tomorrow.*

9. *Infinitive for gerund.* This is the reverse of (8); here an infinitive complement occurs when a gerund is called for.

- a. *I like to eat fish.*
- b. **I enjoy to eat fish.*

10. *Infinitive marker on naked infinitives.* Most verbs that take infinitive complements take the *to*-infinitive; however, a small class of causatives and verbs of perception take the infinitive with no marker, the so-called naked infinitive. Because of the higher frequency of the *to*-infinitive, this error is one of the more common.

- a. *I told Fred to revise the paper.*
- b. *I make/let/had/saw Fred revise the paper./*I made/let/had/saw Fred to revise the paper.*

11. *Tensed that-complement for untensed.* Most verbs which take *that*-complements occur with a tensed verb; however, a few do not. This same pattern holds for certain adjective and noun complements as well.

- a. *I know that he is here every day at noon.*
- b. *I request/require/demand that he be here every day at noon./*I request/require/demand that he is here every day at noon.*

12. *Preposition missing.* A number of verbs in English require a particular preposition or set of prepositions to convey a particular meaning (or in some cases to be grammatical at all). It is possible to analyze errors involving the absence of the prepositions in such cases as subcategorization errors, since they involve placing a verb in a sentence frame containing a missing required category (the preposition). In

most cases the result is a verb-plus-noun phrase in place of the expected verb plus prepositional phrase.

- a. *You can rely on Tom./*You can rely Tom.*
- b. *I listened to the music./I heard the music./*I listened the music.*

13. *Preposition added.* In some cases students add a preposition to a verb that is subcategorized to occur without one.

- a. *I watched the boys play./*I watched at the boys play.*
- b. *I congratulated my friend./*I congratulated to my friend.*

14. *Incorrect preposition.* Given the idiosyncracies of the English prepositional system, one of the more common errors is producing a sentence containing a verb co-occurring with the wrong preposition. Strictly speaking, this is not a subcategorization error since the proper category (preposition or prepositional phrase) is present. Because prepositions belong to a closed set, however, and sometimes indicate primarily grammatical relationships (for example, the passive *by* and the indirect object marker *to*), they are often classified as function words rather than content words and as such particular ones could legitimately be specified in a subcategorization frame. We assume, for the sake of simplicity, that preposition errors of this type involve subcategorization, bearing in mind that in many cases they could also be analyzed as a collocational word choice error or an error within a given idiomatic phrase.

- a. *I am married to Ellen./*I am married with Ellen.*
- b. *I am disappointed with/in you./*I am disappointed of you.*

While the above list of subcategorization errors for verbs is not exhaustive, it covers most of the frequent types of errors and provides enough in the way of representative examples for the discussion that follows. In the next section, we offer some general suggestions for determining whether a given error that we have classified as involving incorrect subcategorization really is such an error in the student's grammar. We then discuss some of the sources of lexical subcategorization errors.

Is It Really a Subcategorization Error?

Students at lower levels, or those at more advanced levels whose understanding of the syntactic structures and rules of English grammar is deficient, may produce certain of the above errors for reasons other than simply having an incorrect subcategorization frame for a given verb. Intermediate students, for example, may have had training and experience with infinitive complements but not understand the English gerund construction well enough to be able to use it productively. For such students an error like that in (9b) (**I enjoy to eat fish.*) may have as its source the lack of control over the gerund construction itself. Similarly, the untensed *that*-complement construction may even be

unfamiliar to rather advanced ESL students, due both to its relative rarity and to the fact that it is not normally introduced until fairly late in the typical structurally based syllabus. An error like that in (11b) (**I request/require/demand that he is here every day at noon*), then, may be based on a general lack of awareness that the untensed *that* complement structure itself even exists. Consequently, before analyzing a given error as having an incorrect subcategorization of the verb as its source, it is necessary first to be certain that the student has sufficient knowledge of the grammatical construction represented in the subcategorization frame.

In some cases, it will be obvious to the teacher from previous experience with the student whether or not he or she is familiar with the grammatical construction in question. If prior experience with the student is not sufficient for making the decision, the teacher may be able to determine whether or not the student controls that construction by looking elsewhere in the current composition or previous ones. Has the student correctly used the construction in other sentences? If so, the error is likely to have an incorrect subcategorization as its source. A further way of making this determination, and the most effective one if the teacher has the opportunity, is to ask the student to attempt to correct the error (after merely indicating its presence but not the correction) and then to discuss the student's response.

Preventing and Correcting Subcategorization Errors: Some General Points

In a real sense, all lexical subcategorization errors stem from an incomplete knowledge of the word in question. There are at least four possible reasons for a student's using a word for which he or she does not have the appropriate subcategorization frame.

1. Lifting an unfamiliar word from a bilingual dictionary. As a general rule, a student's writing will probably be smoother, and there will be fewer errors in both subcategorization and word choice, if the student relies on the words he or she productively commands. If a bilingual dictionary is to be used, it is best to find one that at least gives sufficient examples for the student to be able to infer the common subcategorization frames for the verbs. Better still, the student should be taught to use the bilingual dictionary in tandem with learners' dictionaries (such as those from Oxford and Longman) which give the necessary subcategorization information along with other useful data.

2. Different lexical entries for learned and acquired English systems. Several current models of the language learning process, in particular Krashen (1982), make a distinction between linguistic information that has been learned and linguistic information that has been acquired. *Learned* information is conscious knowledge of the language that is the result of explicit instruction, while *acquired* information is unconscious knowledge resulting from exposure to the target language in

a natural environment. Whether or not one accepts Krashen's sharp division between these two types of knowledge, it is clear that students have information about words that they in some sense "know" but that they somehow fail to use this information in spontaneous speech or writing. For example, a student may have learned that *enjoy* and a number of other verbs take gerund complements, be able to correctly identify *enjoy* as belonging to that class on a test, and even be able to identify the error in (9b) when it is presented in isolation. Nevertheless, the student may still produce an error like (9b) in a composition. If Krashen is right, no amount of overt error correction in this case will shift the proper subcategorization frame of *enjoy* into the student's acquired system for use in spontaneous production. However, since many writing tasks are monitorable, the key to correcting this type of error would be to work on the student's proofreading skills. Students should be given a fair amount of training in looking at their own writing objectively and bringing all of their learned knowledge of English into play in the proofreading process.

3. Interference with the subcategorization of the L1 counterpart. While little is known about the internal structure of the EFL learner's lexicon, the pattern of errors suggests that students sometimes learn the form and meaning of an English word but link its subcategorization to the subcategorization frame of the same or a similar word in the L1. For example, in (13b) (**I congratulated to my friend*), the occurrence of *to* may result if the student's native language marks the noun phrase complement of its word for *congratulate* as an indirect object (e.g., with adative case ending). Similarly, intrusive *be* errors such as those in (7b) (**The strange event was occurred last May*/**The problem is exist today in many countries*) may result from a student's L1 marking intransitives in this class with morphology that either is identical with or overlaps that used in passive constructions. If the teacher is familiar with the student's native language, some reflection on the teacher's part may help in identifying interference as the source of a particular subcategorization error. If not, the teacher can simply ask the student about the behavior of the verb in question. Besides potentially yielding the source of the given error, this technique has the added advantage of increasing the teacher's understanding of the student's native language and showing the student that the teacher has an active interest in it, which may in turn significantly increase the student's motivation to perform well in that particular class.

4. Semantic similarity to another English word with a different subcategorization frame. Another source for subcategorization errors might be called language-internal interference. This type of error occurs when a word which is only partially known is linked to the subcategorization frame of a more familiar one which is semantically similar. For example, *like* and *enjoy* are more or less synonymous in many contexts, and *like* is more frequent and is normally learned

before *enjoy*. As a result, an error like that in (9b) (**I enjoy to eat fish*) may stem from subcategorizing *enjoy* as allowing an infinitive complement by analogy to *like*. The extra *at* in (13b) (**I watched at the boys play*) may similarly have been added by analogy to *look at*. Making the student aware of such errors and having him or her maintain a list of troublesome verbs may decrease the chances of repeating this type of error for a given verb.

Lexical Redundancy Rules and Semantically Based Subcategorization Patterns.

Two final aspects of subcategorization that deserve mention are lexical redundancy rules and semantically based subcategorization patterns. *Lexical redundancy rules* are rules hypothesized by theoretical linguists to account for generalizations in lexical subcategorization frames, many of which were handled by transformations in earlier versions of transformational grammar. As is the case with other types of lexical rules, such as derivational or word formation rules, lexical redundancy rules have exceptions, which is why they have been dropped from the syntactic component in contemporary versions of transformational grammar and most other syntactic theories.

An example of a lexical redundancy rule is the one which captures the generalization that a particle can occur either immediately after the verb or after the object noun phrase, a generalization previously accounted for by the *particle movement* transformation. For instance, a two-word verb like *turn off* which appears in a sentence like *He turned off the light* can also appear in a sentence like *He turned the light off*. This relationship is handled in the lexicon by a lexical redundancy rule which states that any verb which can occur in the structure [v + particle + noun phrase] can also occur in the structure [v + noun phrase + particle], thus eliminating the need to specify the second subcategorization frame individually for each two-word verb. The few exceptions to the rule (such as *run into* in [5b]) are then marked as such in the student's internal lexicon as they are learned, with the verbs fitting the rule left unmarked.

Because they aid students in making generalizations similar to those made by native speakers, there is a certain value in teaching lexical redundancy rules, and formulations of many of them commonly appear as "rules of grammar" in grammar textbooks. It is important to remember, however, that they differ from the purely syntactic rules such as relative clause formation in that they do have exceptions, often a significant number.

Semantically based subcategorization patterns, on the other hand, differ from lexical redundancy rules in that there is a clear semantic basis for their existence, which gives the students a little more to hold onto in learning them. These patterns sometimes underlie lexical redundancy rules but just as often provide a reason for exceptions to those rules. For example, the prohibition on stative verbs occurring in the progressive aspect (as in [4b] **I am knowing French*) is an

example of a pattern that has its basis in the semantic feature of stativity. Another intriguing example of a semantically based subcategorization pattern is the *Bolinger principle* discussed in Celce-Murcia and Larsen-Freeman (1983, pp. 434-436). This principle, named for its originator Dwight Bolinger, is used to distinguish verbs which take gerund complements from those which take infinitives. Briefly, the Bolinger principle states that if the action in the complement of a verb is in progress or already completed during the time frame of the main verb, then the complement will be a gerund, whereas if the action in the complement occurs after the time frame of the main verb or is hypothetical, then the complement will be a *to*-infinitive. For example, in the sentence *I recall getting/*to get up at 5:00 every morning when I had my paper route* the *getting up* occurred prior to the recalling of it so only the gerund can occur. In the sentence *I want to leave/*leaving at noon*, on the other hand, the act of leaving will occur sometime after the act of wanting, so the infinitive is used. For a verb such as *stop*, which is subcategorized to take either type of complement, the principle is clearly illustrated. In the sentence *I stopped drinking*, the drinking is already in progress at the time of the stopping, while in *I stopped to drink (or to take a drink)*, the stopping precedes the drinking. The Bolinger principle can be an aid to students in guessing the subcategorization frame for the complements of less familiar verbs; however, due to the large number of apparent exceptions to it, it should be applied with caution (see Celce-Murcia & Larsen-Freeman [1983, pp. 437-440] for further discussion).

Through the examples and discussion above, we have attempted to clarify the nature of lexical subcategorization errors in order to help EFL writing teachers distinguish them from other types of grammar errors and deal with them at a lexical rather than a syntactic level. Even though all of the error types above have been noted elsewhere in the literature, in such works as Burt and Kiparsky (1972), Richards (1973); and Celce-Murcia and Larsen-Freeman (1983), their relationship to subcategorization has rarely, if ever, been made explicit. Similarly, while the idiosyncratic nature of most of the errors we discussed is probably well-known to the majority of EFL writing teachers, we hope the taxonomy of major error types we have presented and our discussion of their possible sources will aid in dealing with subcategorization errors in a more systematic fashion. If a teacher is going to bring the student's conscious attention to errors for explicit correction, it is useful to have as much information as possible about both the type of error and its underlying cause available for the student. ■

¹ We would like to thank John McVicker and John Graney for helpful comments on an earlier version of this paper and Keiko Tamura for assisting us with the research on subcategorization. Any errors or omissions are our own.

Philip Hubbard is lecturer in linguistics and coordinator of summer programs in English for foreign students at Stanford University.

Donna Hix teaches ESL for the Ohio program of intensive English at Ohio University.

References

Burt, M. K. & Kiparsky, C. (1972) *The gooficon: A repair manual for English*. Rowley, MA: Newbury House.

Celce-Murcia, M. & Larsen-Freeman, D. (1983). *The grammarbook: An ESL/EFL teacher's course*. Rowley, MA: Newbury House.

Chomsky, N. (1957). *Syntactic structures*. The Hague: Mouton and Co.

Chomsky, N. (1965). *Aspects of the theory of syntax*. Cambridge: MIT Press.

Chomsky, N. (1982). *Lectures on government and binding: The Pisa lectures*. Dordrecht: Foris Publications.

Hubbard, P. (1983, October). *Relational grammar and language teaching*. Paper presented at the Midwest Regional TESOL Convention, Minneapolis, MN.

Krashen, L.S. (1982). *Principles and practice in second language acquisition*. Oxford: Pergamon Press.

Radford, A. (1981). *Transformational syntax: A student's guide to Chomsky's extended standard theory*. Cambridge: Cambridge University Press.

Richards, J.C. (1973). A noncontrastive approach to error analysis. In J. Oller & J. Richards (Eds.), *Focus on the learner: Pragmatic perspectives for the language teacher* (pp.96-113). Rowley, MA: Newbury House.