Character Structure of Classic 
STAR WARS 
vs. Its Lesser Prequels. 
A Grammatical Analysis.

Presented at the 33rd Popular Culture Association 
and 25th American Culture Association Annual 
Conference. 
April 19, 2003 

by Alan D. Manning 
Dept. of Linguistics 
Brigham Young University 
Provo, Utah 84602 
alan_manning@byu.edu

This paper will present further refinements to theory 
developed in Manning (1992) and (1996), and most 
recently presented at the 2002 PCA convention as an 
analysis of classic and later versions of Star Trek. 
[Note 1].

Compared to the original trilogy, Star Wars Episodes 
I and II have failed to recapture audience imagination 
in several ways, but this paper focuses on the 
problem of character development. The original 
WARS trilogy maintains a stable ecology of seven 
main (protagonist) characters: Leia, Luke, Han, Obi-
wan, R2, Chewbacca, and 3PO. Darth Vader as the 
major antagonist can be placed outside the core 
"family" group. This paradigm of core characters 
have relational structure in common with characters 
of Classic STAR TREK, GILLIGAN'S ISLAND, 
M*A*S*H, and even BUFFY THE VAMPIRE 
SLAYER. These creations, strong pop icons all, also 
show structural similarity to Shakespeare's most 
familiar play, HAMLET.

Analogies in character pattern among these programs 
point to a generative grammar of popular dramatic 
characters. The grammar predicts whether or not a 
program remains watchable in multiple performances 
to the same audience, the very root of the original 
STAR WARS trilogy's success.

Further central characters introduced in THE 
EMPIRE STRIKES BACK, i.e. Lando Calrissian and 
Yoda, are presented in systematic ways that expand 
but do not disturb the original character pattern. By 
this same process, a structural position is made 
available for Vader in the expanded pattern. This 
position is in place before it is revealed that Vader is 
Luke's father. The same orderly expansion is not 
found in the prequel characters as they are developed 
in Episodes I and II. Grammatical inconsistencies in 
Episode I are only made worse by further 
developments in Episode II.

Let's review what a grammatical analysis is. In most 
basic terms, a rigorous scientific grammar attempts to 
identify the essential parts of a sentence (in whatever 
language). If essential parts of the sentence are 
missing, the sentence fails. If essential parts are 
redundantly duplicated, the sentence also fails.

a. The wookie snarled at the robot. 
b. The wookie did snarl at the robot.

Alternative versions of this sentence are possible 
because the verb "growled" actually has two parts, 
the verb itself and the past-tense element, which can 
be split out and separately represented by the 
auxiliary "did." If tense (present or past) is 
redundantly added back to the verb at this point, the 
sentence breaks down.

c1. *The wookie does snarls at the robot.

c2. *The wookie did snarled at the robot.

Conversely, if the tense element is removed 
altogether, the sentence breaks down.

d. The wookie ø snarl at the robot.

So we conclude that tense is one of the essential parts 
of a well-formed, free-standing sentence 
(=proposition). For the sake of brevity (and people's 
generally low tolerance for long discussions of 
grammar), I'll skip over detailed discussion of the 
other essential parts and simply list them. Linguists 
generally agree that there are seven key components 
of any free-standing proposition in any language 
[Note 2]:

1. Optional modifier-phrases like "at the robot."
2. Subjects like "the wookie."
3. Tense elements like "did."

4. Verbs, like "snarl."
5. The predicate of a sentence, the link between tense and verb: "snarl" or "did + snarl."
6. The expanded predicate of a sentence, the link that ties any optional modifiers to the predicate: "...did snarl at the robot"
7. The main proposition: the link that ties the subject to the predicate (or to the expanded predicate): "The wookie did snarl at the robot."

Two other brief points I need to make about grammatical structure, before discussing what the StarWars-character grammar looks like.

First, there's a reason that there would be exactly seven key elements. The first three elements can be thought of as "atoms." These atoms can either be pure or they can be tied together in four other, logically possible ways. The three atoms of grammar are these three functions:

1. Picture creations = modifiers in their pure form.
2. Pointing actions = subjects in their pure form.
3. Linking relations = connections made between any two parts of a sentence. The essential function of "did" or "does" or "is" or "was" is to tie a subject like "the wookie" to a verb like "growl(ing)."

The other four elements are actually "molecules" made of two elements tied together:

1+2: Verbs are a mix of picture creation and pointing action. The verb "sit" paints a picture but also points us to a "sitter", whatever is doing the action indicated by the verb.

1+3: Predicates are a mix of picture creation and link relation. "Did sit" ties the picture element of the verb (1) to "did" or "does" (3).

2+3: Propositions are a mix of pointing action and link relation. This is illustrated by the fact that very most minimal possible sentence is something like "It did", a subject (2) and a tense element (3).

1+2+3: Expanded predicates. This element ties modifiers like "on the mat" (1) to the rest of the proposition "it did" (2+3).

And my final point about grammar is this. Given the logic of three atoms and four molecular combinations, the seven key elements can be grouped in three distinct but overlapping sets of four elements each:

The first set, metaphorically speaking, are like four variations (a-b-c-d) on the color yellow:

1a. modifiers = pure yellow = picture creation (1)
1b. verbs = orange = yellow (1) + red (2)
1c. predicates = green = yellow (1) + blue (3)
1d. expanded predicates = grey = yellow (1) + red (2) + blue (3)

The second set, metaphorically speaking, are like four variations (a-b-c-d) on the color red:

2a. subjects = pure red = pointing action (2)
2b. verbs = orange = yellow (1) + red (2)
2c. propositions = purple = red (2) + blue (3)
2d. expanded predicates = grey = yellow (1) + red (2) + blue (3)

The third set, metaphorically speaking, are like four variations (a-b-c-d) on the color blue:

3a. tense = pure blue = linking relation (3)
3b. predicates = green = yellow (1) + blue (3)
3c. propositions = purple = red (2) + blue (3)
3d. expanded predicates = grey = yellow (1) + red (2) + blue (3)

To summarize to this point, I've identified three hallmarks of grammatical structure:

1. seven key elements,
2. which are generated by three atomic elements plus four logical combinations of those atoms,
3. which manifest themselves as three distinct but overlapping sets of four elements each.

All of the above brings me to my main theme. The characters in most television programs or film (obviously) would not manifest this kind of
structure—what are the odds that they would? However, there are a small handful of TV programs and film that do manifest such structure, and these share one key attribute in common, which is that they remain endlessly watchable in rerun/repeat performance for a significant number of viewers.

*Gilligan's Island* is perhaps the cleanest example. With other programs the grammatical water is muddied to greater or lesser degree by main-character replacements and by recurring minor characters that appear in a few episodes but disappear over the course of the series. *Gilligan's Island* however makes do with the same seven characters (and even relatively few guest stars) over the whole length of its 96-episode (and subsequent 38-year run on our TV sets and in our heads).

1. Mary Ann
2. The Skipper
3. Ginger Grant
4. Gilligan => 1+2
5. Mrs. Howell => 1+3
6. Mr. Howell => 1+2+3
7. The Professor => 2+3

That the program's characters have grammatical structure is manifest in their endlessly successful rerun cycle, plus the fact that its seven characters transparently sort themselves into three distinct but overlapping sets of four members each:

1a. Mary Ann = pure yellow = Character Known by Name (1)
1b. Gilligan = orange = yellow (1) + red (2)
1c. Lovey Howell = green = yellow (1) + blue (3)
1d. Thurston Howell III = grey = yellow (1) + red (2) + blue (3)

2a. The Skipper = (pure) red = (most) Male character (2)
2b. Gilligan = orange = yellow (1) + red (2)
2c. The Professor = purple = red (2) + blue (3)
2d. Thurston Howell III = grey = yellow (1) + red (2) + blue (3)

3a. Ginger Grant = (pure) blue = (most) Famous Character (3)
3b. Lovey Howell = green = yellow (1) + blue (3)
3c. The Professor = purple = red (2) + blue (3)
3d. Thurston Howell III = grey = yellow (1) + red (2) + blue (3)

Note the exclusion of Ginger from the first set and inclusion of the Professor in the third set. Both these points may require some comment. Ginger Grant is a transparent homage to Marilyn Monroe. We can safely assume then that GG's name is merely a stage name, just as MM's was. Thus, Ginger's true name is unknown, which places her outside of the first "yellow" set. The Professor is identified in the pilot episode as a well-known research scientist and boy scout leader. In comparison, Mary Ann, the Skipper, and Gilligan are utterly obscure.

It seems certain that this clean placement of characters into three logical sets was merely a happy accident. No other program created by Sherwood Schwartz has this precise structure, not even the dismally but mercifully short-lived "Dusty's Trail" (with Bob Denver) which transparently attempted to recreate the Gilligan's Island characters. *The Brady Bunch* comes close perhaps, with its three parallel sets of 1) girl children, 2) boy-children and 3) parents, including Alice as "middle" parent. Each set lacks a fourth member however, and none of the sets overlap.

A fully grammatical version of *The Brady Bunch* would more closely resemble the overlapping structure of *Hamlet* (likewise a story of two families blended, but tragically so). This play is without question the most frequently replayed and requoted of Shakespeare's works:

1a. Ophelia = (pure) yellow = Character Driven (most) Insane by Love (1)
1b. Hamlet = orange = yellow (1) + red (2)
1c. Gertrude = green = yellow (1) + blue (3)
1d. The Ghost Father = grey = yellow (1) + red (2) + blue (3)

2a. Laertes = (pure) red = Character (most) Driven by Need for Revenge (2)
2b. Hamlet = orange = yellow (1) + red (2)
2c. Claudius = purple = red (2) + blue (3)
2d. The Ghost Father = grey = yellow (1) + red (2) + blue (3)

3a. Polonius = (pure) blue = (most) Useless Parent (3)
3b. Gertrude = green = yellow (1) + blue (3)
3c. Claudius = purple = red (2) + blue (3)
3d. The Ghost Father = grey = yellow (1) + red (2) + blue (3)

It's worth noting that from Polonius' family (daughter [1], son [2], and the old windbag himself [3]) constitute the "pure" elements of the Hamlet-grammar. Gertrude and her family (Hamlet, Claudius, The Ghost) constitute the mixed elements:

*Here's the story, of a Pompous Windbag,*
*Who was bringing up his own boy and a girl,*
*One of them was slightly cracked, like her mother,*
*The other one a churl.*

*Here's the story, of a Danish Lady,*
*Who was bringing up one deeply troubled son....*

Coming back to modern television, a very Gilligan's-Island-like pattern is found in *Buffy the Vampire Slayer,* at least in its “classic” form. The program has likewise suffered since Season 3 from actor/character replacements, but this is the subject of another paper. Classic Buffy looks like this:

1a. Willow = (pure) yellow = Special Powers (1)
1b. Buffy = orange = yellow (1) + red (2)
1c. Cordelia = green = yellow (1) + blue (3)
1d. Angel = grey = yellow (1) + red (2) + blue (3)

2a. Xander = (pure) red = Physical Fighters (2)
2b. Buffy = orange = yellow (1) + red (2)
2c. Oz = purple = red (2) + blue (3)
2d. Angel = grey = yellow (1) + red (2) + blue (3)

3a. Giles = (pure) blue = Special Wisdom (3)
3b. Cordelia = green = yellow (1) + blue (3)
3c. Oz = purple = red (2) + blue (3)
3d. Angel = grey = yellow (1) + red (2) + blue (3)

Yet another example is Classic Star Trek, the main subject of my presentation last year. [See also Note 3] It is of the same type as those illustrated above, seven main characters in three overlapping sets. The other 423 people on the classic Enterprise, and the ship itself are (relatively speaking) merely props and background for these seven. In other programs/films with a definitive enemy (like Darth Vader, or Khan in the 2nd Star Trek film) that character defines the “deep” background of “black” as opposed to “blue.”

1a. McCoy = (pure) yellow = Character (most) Prone to Emotional Display (1)
1b. Kirk = orange = yellow (1) + red (2)
1c. Sulu = green = yellow (1) + blue (3)
1d. Scotty = grey = yellow (1) + red (2) + blue (3)

2a. Spock = (pure) red = Character (most) Driven by Concern for Order and Logic (2)
2b. Kirk = orange = yellow (1) + red (2)
2c. Uhura = purple = red (2) + blue (3)
2d. Scotty = grey = yellow (1) + red (2) + blue (3)

3a. Chekov = (pure) blue = Character (most) Defined by Support Function (3)
3b. Sulu = green = yellow (1) + blue (3)
3c. Uhura = purple = red (2) + blue (3)
3d. Scotty = grey = yellow (1) + red (2) + blue (3)

And these third-set "blue" characters are (in turn) merely supportive background for McCoy, Spock, and Kirk. The emotion/logic dichotomy in Classic Trek is well-understood and widely commented on, at least as far as McCoy, Kirk, and Spock are concerned. This emotion-logic dynamic is less noticeable in the background characters, precisely because they are noticed less, but it is still operative. Sulu and Scotty (1-3, and 1-2-3), as few lines as they get, are much more likely to be shown smiling broadly, or making a passing wisecrack, than either Uhura (2-3) or Chekov (3). These latter characters tend to be, like Spock, quite dour in expression. This distinction is in keeping with the overall grammar of the show.

This grammar apparently has considerable psychological impact, and makes Classic Trek what we know it to be. As with *Gilligan's Island* and *Hamlet,* Classic Trek has done exceptionally well in rerun, and it's allusions and turns-of-phrase have utterly permeated popular culture. The same cannot be said for later versions of later versions of Trek.
**AD Manning**

*Voyager*, (which is very like Episode I in structure, as will be shown below) we find too many characters for the logical positions available. As a result, the characters fragment into poorly connected subsets. Like *Next Generation*, the show also has main characters (Kes and Kim) which fill no apparent role at all.

1a. The Doctor = (pure) yellow = Character (most) non-human (1)
1b. Neelix = orange = yellow (1) + red (2)
1c. Tuvok = green = yellow (1) + blue (3)
1d. Seven of Nine = grey = yellow (1) + red (2) + blue (3)

2a. Lt. Torres = (pure) red = Misfit Character (most)
Internally Conflicted (2)
2b. Paris = orange = *not yellow (1), but slightly less red (2)
2c. Chakote = purple = (only somewhat) red (2) + blue (3)
2d. Seven of Nine = grey = yellow (1) + red (2) + blue (3)

3a. Capt. Janeway = (pure) blue = Character (again)
Leading by Building Consensus (3)
3b. Tuvok = green = yellow (1) + blue (3)
3c. Chakote = purple = red (2) + blue (3)
3d. Seven of Nine = grey = yellow (1) + red (2) +
*not really all that blue (3)

Readers may have noticed that the "star" character in a well-formed pattern is always in the orange position: Gilligan, Hamlet, Buffy, and Capt. Kirk, for example, but also Hawkeye on *MASH*, Sam Malone on *Cheers*, and so forth. Data (accidentally I think) ends up the orange character also, and indeed, most NG episodes end up being about him. On *Voyager*, (again surely by accident), Neelix ends up being the true orange character and thus too much the focus of that program, a character too ridiculous in concept and execution to carry the show.

Although it is manifesting a few inconsistencies, there is some sign of grammaticality in Paramount’s latest official Trek offering called *Enterprise*, a 22nd-century prequel to the classic 60s show. For one thing, Paramount somehow realized that it likewise needed to return to the discipline of seven characters, and that it was possible to alter details of the original characters but still keep the overall system intact:

(orange) Scott Bakula as Capt. Jonathan Archer of the NX-01 prototype *Enterprise*.
(red) Connor Trineer as the scrappy southern sidekick, Trip (the chief engineer),
(yellow) Jolene Blalock as the seductive (Vulcan) female, T’Pal.

(green) Anthony Montgomery as the (black) child-prodigy navigator Boomer
(purple) Linda Park as the (Asian) linguist-interpreter Hoshi.
(grey) Dominic Keating as very British tactical officer Reed.
(blue) John Billingsley as the ship's physician, endlessly patient and philosophical Dr. Phlox, who looks VERY like the Dr. Lazarus of *Galaxy Quest*.

But finally we will turn to our main theme, the grammar of Star Wars characters. Classic Star Wars has the same grammatical pattern among its main characters as we have seen:

1a. Leia = (pure) yellow = Character most committed to the rebellion. (1)
1b. Luke = orange = yellow (1) + red (2)
1c. Obiwan = green = yellow (1) + blue (3)
1d. R2D2 = grey = yellow (1) + red (2) + blue (3)

2a. Han = (pure) red = Character (most) reluctantly co-opted into the rebellion (2)
2b. Luke = orange = yellow (1) + red (2)
2c. C3PO = purple = red (2) + blue (3)
2d. R2D2 = grey = yellow (1) + red (2) + blue (3)

3a. Chewbacca = (pure) blue = Character (most)
Defined by Fantasy/Sci-fi Cliche-Support Function (3)
3b. Obiwan = green = yellow (1) + blue (3)
3c. C3PO = purple = red (2) + blue (3)
3d. R2D2 = grey = yellow (1) + red (2) + blue (3)

And this pattern, I argue, is actually the driving force behind the original popularity and endless watchability of the original series. When further characters are added in Empire Strikes Back that
continue in Return of the Jedi, they are introduced in a way that does not disturb the original pattern in a significant way. What happens is that new characters are effectively pair bonded with the older, pre-established grammatical characters.

Yoda is bonded to Obiwan; they work as a team. Chewbacca is bonded to C3P0; they are literally one creature in the climax of Empire. Han is put into carbon freeze and effectively replaced at the controls of the Millenium Falcon by Lando.

In in terms of the color pattern we’ve been describing, two separate but closely related values for each of the secondary colors are now allowed, two shades of green, one more yellow and one more blue, two shades of purple, one more blue and one more red, and two shades of orange, one more yellow and one more red. The two orange characters in classic Star Wars are Luke and Han, now equal co-stars and both equally the center of attention in the film.

1a. Leia = (pure) yellow = Character most committed to the rebellion. (1)
1b. Luke/Han = orange = yellow (1) + red (2)
1c. Obiwan/Yoda = green = yellow (1) + blue (3)
1d. R2D2 = grey = yellow (1) + red (2) + blue (3)

2a. Lando = (pure) red = Character now (most) reluctantly co-opted into the rebellion (2)
2b. Luke/Han = orange = yellow (1) + red (2)
2c. C3PO/Chewbacca = purple = red (2) + blue (3)
2d. R2D2 = grey = yellow (1) + red (2) + blue (3)

3a. DARTH VADER = (pure) blue = Character (most) Defined by Fantasy/Sci-fi Cliche-Support Function (3)
3b. Obiwan/Yoda = green = yellow (1) + blue (3)
3c. C3PO/Chewbacca = purple = red (2) + blue (3)
3d. R2D2 = grey = yellow (1) + red (2) + blue (3)

Note that at exactly this point, room is made in among the protagonist characters for Vader, at precisely the point where we realize that he was Luke’s father and thus once a hero/goodguy himself, like the other focal characters. In turn, the Emperor is introduced and made the new “black” character in the far background.

So what goes wrong in the second trilogy of films? In many ways the ungrammatical properties of Episode I resemble the problems of Star Trek: Voyager. There are too many characters who are not key to the narrative, not clearly slotted in the grammar, but who still get significant screen time: Anakin’s mother, Watto, C3PO and Mace Windu for instance.

Not all characters effectively fit their approximate grammatical slots. Jar-Jar Binks is of course the most glaring example. A true warrior figure is called for in the grammatical pattern that ought to be present, based on the definitive overlapping features of the existing characters. In other words, Qui-Gon, R2, and Amidala are blooded vetrans and serious fighters, and this defines the second set. Jar-Jar’s Gungan civilization also is, but Jar Jar himself is not, and this is a catastrophic error. The characters should have been played seriously, a Klingon-like being who falls in with Qui-Gon accidentally and becomes his main support (like Han to Luke) rather than his main headache.

As a further complication, Padme and Queen Amidala occupy different slots in the grammar, and yet it turns out in the end that they are one character.

But the most damaging difficulty of all is that four of the secondary-color characters are not properly bound together as units. Anakin and Qui-Gon perhaps are, but Obi-Wan and Yoda are not, and R2D2 and NuteGunray are not.

1a. Padme = (pure) yellow = Young Assistant. (1)
1b. Anakin = yellow-orange
1c. Obiwan= yellow-green
1d. Amidala = grey = yellow (1) + red (2) + blue (3)

2a. Jar Jar = (pure) red = Displaced Warrior (2)
2b. Qui-Gon = orange-red
2c. R2D2 = red-purple
2d. Amidala = grey = yellow (1) + red (2) + blue (3)

3a. Palpatine = (pure) blue = LEADER (3)
3b. Yoda = green-blue
3c. Nute Gunray = blue-purple
3d. Amidala = grey = yellow (1) + red (2) + blue (3)
It’s fairly apparent that Lucas was heavily influenced by criticism of Episode I as he retooled Episode II. Jar-Jar is made a minor character, but also the crowd of central characters in the secondary colors is cleaned out. The problem is that Lucas went too far, and the characters that remain are effectively separated from each other for the first two thirds of the movie. The film is in effect three movies, defined by the three corner colors of the fragmentary grammar. This result is effectively explained by the removal of secondary characters--Qui Gon is dead; R2D2 is relegated with C3PO to the role of prop and comic relief.

Anakin and Amidala are pushed off by themselves into a rather dull and one-dimensional love story (YELLOW).

ObiWan is moved to the RED position, sent off by himself to hunt down Bobba Fett, in a red spaceship with a red, characterless droid.

Only the last movie, the BLUE movie which is the last half hour, is of any real interest. Here all the clone masters (Yoda, Palpatine, Dooku, and Jengo Fett himself) meet to do battle, and Anakin and Amidala are reunited with Obiwan, to make the three characters which are the center of interest. Mace finally comes in at purple (with a purple lightsaber to boot) but he is very late too the party, too late to give the Episode II characters the cohesion they so badly needed in Acts 1 and 2. Bottom line: Episode II is grammatical at the end, but not for most of the movie:

(Final Act of Episode II)

1a. Padme = (pure) yellow = Doomed Love Characters (1)
1b. Skywalker = orange = yellow (1) + red (2)
1c. Yoda = green = yellow (1) + blue (3)
(has to fight his lost, beloved student Dooku).
1d. Jengo Fett (&beloved son) = grey = yellow (1) + red (2) + blue (3)

2a. Obiwan = (pure) red = Character sent out to track down the master assassin (Jengo Fett).
2b. Skywalker = orange = yellow (1) + red (2)

(tried and failed early in film)
2c. Mace Windu = purple = red (2) + blue (3)
(finally kills Jengo, late in film)
2d. Jengo Fett = grey = yellow (1) + red (2) + blue (3)

3a. Palpatine = (pure) blue = Clone Puppetmaster (3)
3b. Yoda = green = yellow (1) + blue (3)
3c. Mace = purple = red (2) + blue (3)
(bring the clones to battle)
3d. Jengo = grey = yellow (1) + red (2) + blue (3)
(Father of the race of clones).

Bottom line: this model of character grammar has a high degree of explanatory power for problems of popular culture, explaining the success or failure of key icons in both television as well as film.

ENDNOTES

[Note 2] For further discussion of the structure of grammatical phrases and principles generally understood by linguists to apply to all languages, see Steven Pinker, The Language Instinct, NY: William Morrow, 1994, pp 74-118.

[Note 3] The grammaticality of a number of other programs (e.g. M*A*S*H, Cheers, The Beverly Hillbillies, Night Court, Hogan's Heroes, etc.) are discussed in more detail in Manning, "Paralinguistic Character Structure in Popular Syndicated Television." Semiotica (1992), vol. 89, 47-82.

Also, the theory of character patterns is fleshed out in narrative form in Manning, Supposition Error: a novel, Parlay Press, 1996.