Syntactic complexity and coordination in a verbal production task
Preliminary summary of experiment results
Rice Complexity Symposium, Precirc

Marjorie Barker and Eric Pederson
University of Oregon

Abstract

The origin of syntactic complexity is not completely clear. Some degree of syntactic complexity can be seen as the natural consequence of the evolution of a rich communication system, while much cross-linguistic variation must be attributed to historical circumstance with often unclear causal factors. This study addresses the synchronic issue of why speakers elect to employ greater or lesser syntactic complexity for expressive purposes.

We examine the extent to which changing the communicative intent of the speaker affects the degree and type of reliance on syntactic complexity. Participants viewed complex human action video stimuli and were asked to respond in detail to a single question concerning either *what* had happened in the scenario they had just watched or *why* a particular event in the video had occurred. Our prediction was that responses to the *why* question would have more syntactically complex constructions than responses to the *what* question. The experimental results with these stimuli did not straightforwardly uphold the hypothesis; however, there was significant difference in the amount of coordination within intonation units between the two conditions and the types of complementation varied between conditions.

Background

How to describe complexity

One option for defining linguistic complexity is in terms of processing load: utterances that are difficult to process (either in production and comprehension) are by definition complex. Language offers speakers some options in the complexity of the utterances which they produce, although this will be constrained by certain choices of verbs and other reference needs. To the extent that speakers can choose for greater or less complex constructions, this choice may be influenced by other task demands on the attentional and processing systems.

Informational content, how much data is entailed in a given phrase, is another factor in processing. In this regard, pronouns—or even ellipsis—could be considered highly complex, yet it seems that these grammatical elements of language actually facilitate processing, by making it possible to reference words or whole phrases that are highly salient / available in memory.

It could be argued that even a single-word utterance can be complex, in that its intended and perceived meaning involves predication (child-caretaker interaction clearly exemplifies this possibility, as for example when a child says “milk,” and the caretaker understands the utterance as a request and responds “Oh, do you want some milk?”). On this argument, all language is complex. But there are different kinds of complexity. The child’s utterance is meaningfully complex, while the caretaker’s response is more

complex in form. Descriptive linguistics recognizes a difference between simple utterances and those that are more concatenated in form, and again between concatenation and subordination. Complexity is thus relative: a two-word utterance is in a structural sense more complex than one word; conjoined phrases are more complex than a single phrase; one phrase embedded within another is still more complex.

The complexity we seek to explain is structural. Given that varying degrees of syntactic complexity can be used to represent the same event, a logical question follows as to the motivation for using more complex syntax. What causes speakers to use embedded rather than concatenated structures?

The origin of complexity

Evolutionary accounts offer explanations of the development of syntactic complexity in terms of adaptive benefit: slight processing gains can deliver a substantial advantage for quick decision-making and consequently for survival. Developmental accounts describe how verbs representing separate but frequently-connected events may move through stages of paratactic association (coordination) to syntactic complexity (subordination) to complex verb forms like complements. Grammaticalization theory accounts for the various steps in the process of morphosyntactic change, but there is no explanation for why some languages follow the predicted path while other languages do not.

Either of these explanations must remain hypothetical since we cannot gather data on language production and comprehension outside of the laboratory. Lacking from both evolution and diachrony is any data on how language is used online. Attempting to explain how complexity varies online today will help to build a model of the origin of complexity in general, since the synchronic choices behind evolutionary and diachronic change are related to online expressive demands. Therefore we start with an exploration of online language use in this study.

We investigated the conditions under which people use complex syntax (i.e., the hierarchical structures of complementation or relativization) as opposed to simple syntax (i.e., the conjoined structures of coordination, or separate clauses) to describe associated events.

Our initial intuition was that empirical characteristics of the events themselves might affect the choice: on this account, events that are closer in time or space, or events that have a high degree of shared referents, would be more likely to be described with complex clauses. An immediate difficulty arises in this regard, however, because for many common complement constructions it is difficult to develop stimuli that clearly represent separate component events—for example cognition/utterance complements like “he said the train was late”, manipulation complements such as “she had him call a cab”, and modality complements in the vein of “they wanted to arrive early”. It is difficult to imagine, using the last example for instance, how “wanting” could be presented separately from “arriving” in an experiment using visual stimuli. Thus there is problem of lexically determined complementation; in order to describe some situations, complement structures are the only linguistic option, at least in English. Our hypothesis was that these complement-taking verbs would be more used in explanations than in descriptions.

Clearly certain genres typically demonstrate greater syntactic complexity. For example, formal written texts tend to have far more subordination than spontaneous

speech. However this reflects editing processes which are less subject to processing constraints. Since we are interested in the factors driving greater complexity in production, we focus here on the variable conditions of extemporaneous speech.

Motivations for complexity in oral narrative

We presumed that a pervasive function of complementation is evaluative. In the case of manipulation and modality complements, speakers use the main verb of the complement to explain the motivation for the action in the complement. With cognition utterance verbs, speakers express a basis for confidence in the subordinated event. We might further propose that such evaluative information will automatically be packaged this way—it may in fact be difficult or awkward to separate such information syntactically.

Because of this evaluative use of complementation, we hypothesized that speakers would use complements more often in describing a scenario if they were asked to give reasons and motivations rather than to simply give an account of events. In other words, we are looking for online motivations for speakers to increase their syntactic complexity. To test the hypothesis, we decided to present visual stimuli in the form of video clips and ask experimental participants to alternatively "describe" or "explain" what had happened in the videos they viewed.

Experiment 1: Within-subjects describe/explain

Methodology

Ten video clips, each approximately one minute in length, were taken from separate episodes of the USA Network TV series Monk. Segments were chosen that depicted a sequence of events telling a brief coherent story in which two or three main characters were involved (brief descriptions of the ten video clips are included in Appendix A).

Fifteen experimental participants (12 female, 3 male; median age 19) were recruited from the University of Oregon Psychology & Linguistics Human Subjects pool. Participants watched the ten video clips, each approximately one minute in length. Half of the participants were given the following instructions:

> You are going to watch ten video clips selected from the TV series "Monk." For the first five, after you watch each one, I'd like you to describe from memory the events as you saw them happen on the screen, so that someone listening to your description would be able to describe the same events without seeing the video.
>
> For the second set of five, please give from memory an explanation of why the events happened the way they did, so that someone listening to your description would be able to understand how the various actions are related to each other and why the things that happened on the video took place in that way.

The other half of the participants received the same instructions, with the order reversed; i.e., these participants were asked to explain the reasons for the events in the

first set of five videos they viewed, and describe the events in the second set of five videos.

Participants viewed one video clip at a time, and then were tape-recorded as they gave an oral description or explanation of the events in the video. For the ten videos they viewed, each participant produced five "descriptions" and five "explanations." Video clips were presented in a fixed order. The fifteen participants produced ten narratives each for a total of 149 data points (75 "describe", 74 "explain"; one "explain" data point was not recorded due to equipment error).

Results

Results showed no difference between the two conditions in amount of complements, coordination, or relative clauses. The length of the narratives also did not differ between the two conditions.

				length	
	comp.	coord.	rel.	seconds	lines
describe	4.8 (.12/sec) SD=3.9	2.9 (.07/sec) SD=2.8	1.7 (.04/sec) SD=1.4	40.3	7.5
explain	4.6 (.11/sec) SD=3.5	2.3 (.05/sec) SD=2.3	1.6 (.04/sec) SD=1.7	42	7.4
describe + explain	4.7 (.11/sec)	2.6 (.06/sec)	1.64 (.04/sec)	41	7.4

Table 1: Results of Experiment 1 – within subjects describe/explain

In Experiment 1, fifteen participants were asked to either "describe in detail what they had seen" or alternatively to "explain why the events happened the way they did". The prediction was that the explanatory recalls would have greater overall syntactic complexity (e.g. higher subordination: coordination) than the more purely descriptive recalls. Data analysis showed no difference in rates of complementation, coordination, or relativization between the "describe" and "explain" conditions; participants simply used more adverbial "because" clauses in the "explain" condition.

Experiment 2: between-subjects *why* vs. *what*

For the second experiment we decided that rather than asking participants to describe or explain the videos, we would ask them questions which prompted for detailed descriptions or explanations.

We hypothesized that asking questions about why a particular action was taken would point participants toward explanation, because speakers would talk about motivations and causality, requiring the use of complement-taking verbs. On the other hand, we predicted that asking about what happened would not require as many complement structures, since speakers would be talking about concrete visible actions and presumably not about the reasons the events took place. Two of the ten stimuli used in Experiment 1 were eliminated, and eight of the original ten videos were used for the experimental study.

Why condition

Fifteen experimental participants (7 female, 8 male; median age 21) were recruited from the University of Oregon Psychology & Linguistics Human Subjects pool. Participants first watched a sample video clip, then were given the following instructions:

> You will be watching eight more video clips selected from the same TV series. After you watch each one, I will ask you a question about why something happened in the video. Please give a full and complete explanation in answer to the question, in other words, several sentences, connections between various events, rather than simply the immediate cause. For example, for the sample video, if the question was "Why does Monk use a handkerchief to hold onto Kevin's pants?" you wouldn't want to say it's just because Monk has OCD. You need to talk about the multiple things that are going on, such as
>
> > Kevin's need to look into the house
> > the need not to touch the floor
> > the need for a counterweight
> > Monk's fear of germs
> > etc.

Participants viewed one video clip at a time, then were asked one question about the video they had just viewed, and tape-recorded as they produced an oral response to the question.

What condition

Fifteen experimental participants (9 female, 6 male; median age 19) were recruited from the University of Oregon Psychology & Linguistics Human Subjects pool. Participants first watched a sample video clip, then were given the following instructions:

> You will be watching eight more video clips selected from the same TV series. After you watch each one, I will ask you a question about what happened in the video. Please give a full and complete description in answer to the question, in other words, several sentences rather than simply an undetailed response. For example, for the sample video, if the question was "What does Kevin do in the process of getting the pencil?" you wouldn't want to say just that he gets it with the shovel. You need to talk about the multiple things that you saw, such as
>
> > Kevin's having Monk hold his pocket
> > his leaning into the house
> > reaching for the chair
> > getting the shovel
> > using the shovel to scoop up the pencil
> > etc.

Participants viewed one video clip at a time, then were asked one question about the video they had just viewed, and tape-recorded as they produced an oral response to the question (*why* and *what* questions are included with the video descriptions in Appendix A).

Results

For each speech passage produced, we determined the length of the response in seconds. We also counted the number of transcribed lines for each response as a secondary comparative measure in case rate of speech differed greatly between participants. For each response, coordinate, complement, and relative clauses were marked and counted.

Speech passages were divided into phrasal units as the basic processing element. These are considered an indication of planned speech units, hence mental association, in addition to and possibly more reliable than lexical than lexical coordinators like *and, but,* and *or,* which are often used as discourse markers to show narrative continuation rather than syntactic connection. In other words, without knowing the intonational packaging it is impossible to tell from transcripts whether *and* marks coordinate clauses, since coordination joins syntactically complete expressions. Complementation constructions, on the other hand, can span across intonation contours, because the first part is clearly incomplete and awaits the completion in the second. There may well be long-term planning across boundaries, but this is difficult to measure and not necessarily consistent.

Therefore, coordination was described as clauses connected by *and, but, or,* or *Ø.* but verbs were counted as coordinated only if they fell under the same intonation contour. Complementation was determined syntactically – in other words, even if the complement construction occurred across the boundary between intonation contours. Relativization was also determined syntactically.

We did not count clauses from meta-cognition or commentary– i.e., "I guess he was having lunch" does not count as complex for purposes of describing/explaining the action of the video. "Monk thought there was another snake" does count as complex, because it describes what is going on in the video. The verb *like* used for quotation was counted as cognition/utterance complementation.

Complement types

We further analyzed complement clauses by classifying them into six groups:

1. **Speech**—describing speech or semiotics (examples include *she yells at Monk to help, he pointed out that he saw she's saying 'he's gonna get me!'.*
2. **Cognition**—describing mental activity (examples: *he knew how to do it, he couldn't figure out how to move it.*
3. **Manipulation**—describing the use of someone or something to accomplish a task (examples: *he got the ferris wheel to move, he used his legs to press the door shut.*
4. **Modality**—describing attempt, intent, obligation, ability, or possibility (examples: *he tried to climb on, he wants to know.*
5. **Aspect**—describing inception, termination, continuation, success, or failure (examples: *it ends up breaking, he started to back out.*
6. **Other**—where the subordinated clause describes intent (examples: *he climbed up to prevent the guy from getting her, he reached to grab the cloth*.

Fifteen participants who answered one *why* question after viewing each video produced eight narratives each for a total of 120 data points. Fifteen participants who answered one *what* question after viewing each video produced eight narratives each for a total of 120 data points. Results are shown in Figure 2.

why questions n=120 productions (15 participants, 8 videos)

	comp.	coord.	rel.	sec	lines
total	286	86	96	2766	481
mean	2.38	.72	.80	23	4
	SD=1.56	SD=0.85	SD=0.94	SD=7.93	SD=1.43

what questions n=120 productions (15 participants, 8 videos)

	comp.	coord.	rel.	sec	lines
total	305	158	90	2932	460
mean	2.54	1.32	.75	24.4	3.8
	SD=2.21	SD=1.22	SD=1.00	SD=12.1	SD=1.85

Table 2: Results of second experiment – between subjects *why* vs. *what*

One-tailed T-tests were conducted on the complementation and relativization data, since the difference between conditions was in the predicted direction (more complex syntax in the *why* condition). A two-tailed T-test was conducted on the coordination data, since difference between conditions was not predicted. Anovas were run for all three measures comparing *why* vs. *what* responses. These analytic measures showed a slightly higher rate in the *why* condition of complementation (T-test P=.41, single factor ANOVA P=.8) and relativization (T-test P=.14, single factor ANOVA P=.5), while coordination was greater in the *what* condition, with the difference being highly significant (T-test P=.001, single factor ANOVA P<.001).

Although the overall frequency of complements was not significantly different between conditions, the types of complements used by speakers varied noticeably between the *what* and *why* conditions. See Figure 1.

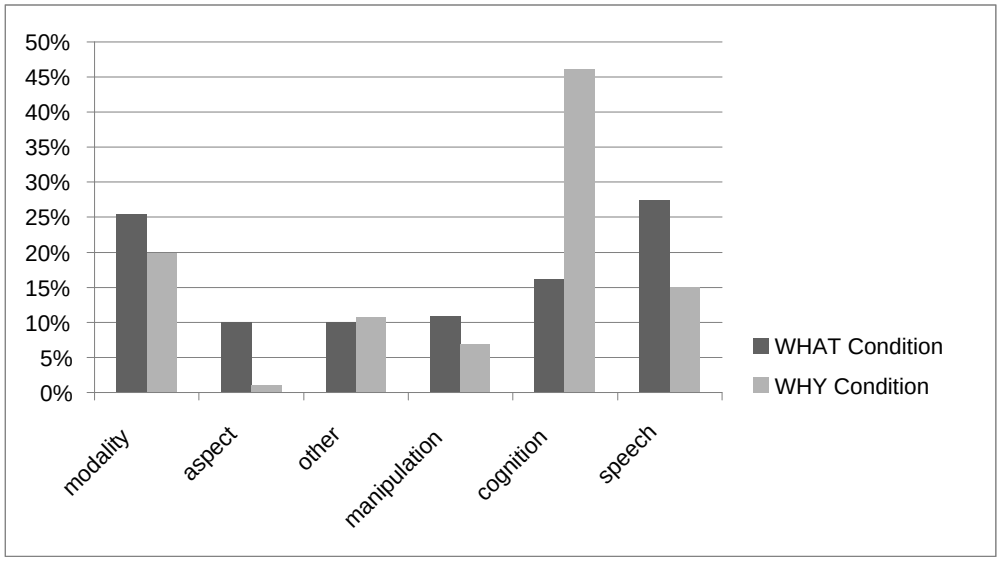

Figure 1: Comparison of complementation types in Experiment 2

Discussion

Our hypothesis that people will use more complex syntax when asked to explain an event than when simply describing was not supported, and thus must be rejected for tasks such as this experiment. The rate of complementation in online speech appears, at least in these tasks, to be fairly consistent. However, the use of complementation does vary with expressive purposes. Participants answering why questions are far more likely to use complements of cognition than those answering what questions.

There was no cognitive loading beyond the task of giving descriptions / explanations; participants were just talking, concentrating only on what they said. Reporting observed events is a common, low-effort task. Based on these results, we might propose that the rate of complementation is stable in speech, possibly because of cognitive constraints on how much embedded information can be processed in a given unit of time. In other words, people typically embed at a rate near their current production capacity. It may well be that in (frequent) situations where the language production system is sharing cognitive resources with other processes that greater differentiation of complexity would emerge.

Surprisingly, however, the amount of coordination varied between conditions in Experiment 2, without affecting the rate of production of complements and relative clauses. This increase of coordination effectively places more clauses within the phrasal unit. Multi-clausal intonation contours in the *what* condition could be interpreted as reflecting a conceptual grouping of events.

A notable difference between the two experiments was the length of response given by participants. In the first experiment, where they were asked to either describe or explain the events that took place in the video, the average response length was 41 seconds. In the second experiment, where participants responded to a specific question about the video, it was 23 seconds, despite specific instructions to answer at length. It seems that participants interpreted Experiment 1 to require a narrative, and Experiment 2 to require a brief and fairly specific answer; as if the first were open-ended and the second a test question. The rate of complementation and relativization was fairly stable between the two experiments despite the large difference in response length. The rate of coordination, however, was more variable, being similar in Experiment 1 and the *what* condition of Experiment 2, while being much lower in the *why* condition of Experiment 2.

Response length

	Ave. response length		**Ave. no. of associated clauses per response**		
	Sec.	**Lines**	**Complement**	**Coordinate**	**Relative**
Exp1 overall	23	4	2.4 (.11/sec)	0.3 (.014/sec)	0.5 (.024/sec)
Exp2 overall	41	7.44	4.7 (.11/sec)	2.6 (.06/sec)	1.64 (.04/sec)

Complex structures

(means)	**P** n=15 **describe**	**P** n=15 **explain**	**P** n=30 **total**	**E** n=15 ***why***	**E** n=15 ***what***
comp./sec	.12 SD=.060	.11 SD=.058	.11	.106 SD=.065	.103 SD=.083
coord./sec	.07 SD=.052	.05 SD=.046	.06	.029 SD=.035	.055 SD=.048
rel./sec	.04 SD=.041	.04 SD=.035	.04	.035 SD=.044	.031 SD=.042

Table 3: Comparison of Experiments 1 and 2

Appendix A: Description of stimuli, *why* and *what* questions

Ball—Monk and a woman are in the office of a sports agent, who is talking on the phone. Monk takes out a tissue and wipes a basketball in a display case. The agent gets off the phone quickly and tells Monk to stop, explaining the "stain" on the ball is Michael Jordan's autograph. Monk proposes that Michael could sign the ball again and the agent sarcastically agrees; Monk says he is relieved and the woman grimaces.

> Why does the woman look disgusted when Monk says he is so relieved?
> What does Monk do with the tissue?

Car—Two men hand Monk the keys to a car parked at a curb. Monk takes the keys and goes to the driver's side of the car, but climbs first into the back seat of the car, then into the driver's seat, while the woman with him gets in the passenger seat. He maneuvers with difficulty out of the parking space while other drivers honk. There is a crash; the car has crashed into a light pole. Monk tells the woman that she told him to turn, so he turned.

> Why did Monk crash into the light pole at the end?
> What did Monk do with the car?

Dog—Monk is running through the aisles of a store that is closed. A Doberman is chasing him. He goes into a customer service booth, lies on the floor, and holds the swinging door closed with his feet. He looks around, grabs a phone, and dials. The scene cuts to a restaurant where a man and woman are eating. The woman's phone rings; she answers but there is no response, only barking. The man listens and then says "let's go." They hurriedly leave the restaurant.

> Why did the man and woman leave the restaurant in a hurry?
> What did Monk do to escape the dog?

Ferris wheel—A woman is in the seat of a ferris wheel; a man is climbing toward her on the bars of the ride. Monk is at the controls of the ferris wheel, pushing buttons. The woman shouts at Monk to get her down. The wheel starts moving, Monk pulls a lever out of the controls, and finally jumps up onto the ferris wheel.

> Why does Monk jump onto the ferris wheel?
> What does Monk do at the ferris wheel?

Leaning—Monk and another man stand in the doorway of a house. They open the door and nobody is home. The other man asks Monk to hold the pocket of his pants and act as a counterweight so he can lean into the house and look around without touching the floor. He can't reach a pencil on the sofa, so grabs a shovel outside the door and uses it to get the pencil. Monk pulls him back. (used as example for experiment)

Sandwich—Monk is in the break room of a store. He takes a bag lunch from a refrigerator. Two coworkers are trying to get candy out of a vending machine. A woman comes in and tells the two to get back to work. The two others make some comments about the woman to Monk. He opens his sandwich and finds that a bite has been taken out of it, and he throws it down. (not used in experiment)

Roof—Monk and another man look out an upper-story window; Monk points out a red rag on the chimney of the building. The other man climbs out onto the roof with Monk cautioning him to be careful. The man gets the rag, and Monk asks him to look for footprints, which he sees, and then climbs back to the room where he says that the rag is evidence.

> Why are Monk and Manny looking for footprints on the roof?
> What did Manny do on the roof?

Room—Monk is in a nursing home room. He examines the furniture, walls, etc and questions a nurse about the occupant of the room and whether anything has been moved. A police officer is talking on a cell phone in the background describing what Monk is doing. He asks Monk to give an opinion, and Monk states that the man was murdered. The police officer seems surprised.

> Why was the captain surprised that Monk said the old man was murdered?
> What did Monk look at while investigating the old man's room?

Snake—Monk and a police officer are inside a house. The officer is putting a snake back into a cage while Monk acts panicky. When the snake is back and the lid closed, Monk examines the snakes, saying there is a feeding schedule for Curly, Larry, and Moe. He looks closely, shouts, then climbs onto the table. The officer runs into the room and asks what he's doing. Monk replies that there are only two snakes in the cage; one must be loose.

> Why did Monk jump up on the table?
> What did the captain do with the snakes?

Tie—A woman is talking to a man on a park bench, while Monk leans over them. The man says they have no proof; Monk stands up and holding his tie, speaks toward it, saying "come and get him." There is no response. Monk runs toward a white van parked nearby and opens the back doors. He asks the policemen inside why they didn't arrest the man; they respond that they heard nothing, and ask what happened to a stain on his tie. He responds that he finally got it out; they look dismayed.

> Why were the agents inside the van upset?
> What did you see Monk do with his tie?

Appendix B: Sample responses

What Condition: ***What did Monk do with the car?***

B16—Alright, it starts out with Monk taking, a pair of keys from a man, on a sidewalk, / and then he walks around the front of the car, / and he sits in the back seat, / and then closes it and opens the door again I believe, and, goes to the front seat where he's driving, / and he sticks his head out the window, / and, tries to back up and hits a car and it makes a noise, / but he drives away anyway and sticks his hand out the window to kinda wave people away, / they're honking, and then, / uh, he hits, uh, I think it's a light post, on, the right side of the car. / And gets out and is very frustrated. (38 sec)

B17—Monk got in the wrong door of the car, in the back seat, instead of the front seat; / he, then, reversed into another car, / and then kept going, / and the alarm was going off, / then he, swerved in, to the middle of the street, while another car was coming, and they honked at him, / and then he kept going, / and then, ran into, a pole. (25 sec)

B18—Monk got in the back seat first cause he was distracted, / and then he got into the driver's seat, / and he, backed up, / and, since he was nervous, it seems like he, wasn't paying attention and he backed up into a car behind him, / and then, continued to drive off, and cut off traffic, / and then, he drove into a pole. (24 sec)

B19—So, Monk got the keys from the guys standing on the curb, / and first he went to the back seat instead of getting into the driver's seat, / and completely closed the door and then got into the driver's seat, / and, s- proceeded to try to pull into traffic, it took him a few minutes, / he looked backwards, / and it sounded like he kinda backed into a car but I'm not really sure if he actually did, / and then he pulled out into traffic and cut somebody off because they honked the horn, / and then, he, ended up crashing into a light pole. (32 sec)

Why condition: ***Why did Monk crash into the light pole at the end?***

B1—It appears that Monk, does not know how to drive, or he wouldn't have gotten into the back seat of the car to begin with, / and in pulling out, he was very jerky, and, as he said to the woman, um, that she had told him to make the turn. / And so he followed her directions. (21 sec)

B2—Monk crashed into the light pole because he was so frazzled from what was going on, / he was trying to play it really cool, / um, getting out of that parking spot, cause he thought it was a really good spot, / and, um, he got stressed out, / and, that girl told him to turn, and supposedly he turned, / and, he was just listening to other people I guess. (23 sec)

B3—He crashed into the light pole because, uh, he was listening to the lady's instructions, / and it didn't look like he had, ever like driven a car before, / because, first when he gets into the wrong seat and, gets in the back seat of the car, instead of the driver's seat; / so, uh, it looked like he had little experience of driving cars, / so, uh, he didn't know when to turn right. (23 sec)

B4—Uh Monk crashes into the light pole – light, pole, at the end, because he, uh doesn't seem to be able to drive very well or be very sure of himself. (12 sec)