THE INTERPRETATION AND REALIZATION OF FOCUS:
AN EXPERIMENTAL INVESTIGATION OF FOCUS
IN ENGLISH AND HUNGARIAN

Beth Ann Hockey

A DISSERTATION
in
Linguistics

Presented to the Faculties of the University of Pennsylvania in Partial
Fulfillment of the Requirements for the Degree of Doctor of Philosophy

1998

______________________________
Supervisor of Dissertation

______________________________
Graduate Group Chairperson
Chapter 1

Introduction

‘When I use a word,’ Humpty Dumpty said in rather a scornful tone, ‘it means just what I choose it to mean -- neither more nor less.’

‘The question is,’ said Alice, ‘whether you can make words mean so many different things.’ (Carroll 1941; Carroll 1993)

Linguists have associated the word “focus” with a wide variety of phenomena. In addition a wealth of other terms including “new,” “emphasis,” “stress,” “rheme,” “comment,” “accented,” “prominent,” “stress,” “informative” and “contrast” have been attached singly or in combination to phenomena that seem to be the same as, similar to or overlapping with those that have been called focus. To get a feel for “focusesque” phenomena consider the difference between the answers in (2) and (3) to the question in (1).

(1) What do aardvarks eat?

(2) Aardvarks eat ANTS.

(3) ARDVARKS eat ants.

Assume for the moment that capital letters mark an item as being spoken in some way that makes it more salient than non-capitalized words. It seems intuitively clear that (2) is a better answer to (1) than (3) is. Having ARDVARKS stand out makes (3) strange or inappropriate as an answer to (1). On the other hand, (3) would be a fine answer to “What eats ants.” This difference between (2) and (3) is generally taken to be a difference between
ANTS being the focus in (2) and ARDVARKS being the focus in (3). There are two aspects to this focus. One aspect is being spoken so as to be more salient. This salience is the way focus is expressed, its realization. The other aspect is the interpretive difference that accompanies differences in focus and makes (2) a better answer to (1) than (3) is. This aspect is the interpretation of focus.

One idea that has been proposed about the interpretation of focus (discussed in more detail in Section 2.2.2.2.2) is that the focus must be new, that is, unknown to the hearer or novel in the discourse. This might seem reasonable for an exchange consisting of (1) and (2). However, (2) could also be the answer to (4).

(4) Do aardvarks eat ants or termites?

When (2) is used as an answer to (4), ANTS has already been mentioned in the question. The hearer already knows about ANTS in the context of this question answer exchange. Therefore, the focus ANTS, cannot be considered new.

In other cases of focus, the alternatives to the focus item seem to be excluded, or alternatively the focus seems to be exhaustive. This exclusion/exhaustiveness effect seems especially strong in it-clefts. It is strange to utter (6) immediately following (5) because we have the sense that termites were eliminated as possible aardvark food by (5).

(5) It is ANTS that aardvarks eat.

(6) #And aardvarks eat termites.
Notice however, that uttering (6) after (2) is felicitous, so again we see some evidence that it might not make sense to interpret all cases of focus in the same way. There are many other sorts of focus phenomena that could be discussed, but the major point would remain the same: the interpretation of focus does not appear to be uniform. The use of the term focus for phenomena with different interpretations is one source of confusion in the literature.

In addition to representing multiple interpretations, the term focus has also been used to designate prosodic prominence (which I have been describing as being more salient), thereby providing another source of confusion. Even when it is clear that interpretation is being discussed, the type of interpretation may remain a mystery. It is not unusual for focus to be treated as though it were a well-defined concept not requiring discussion. The fact that so much work on focus has studied English has probably also contributed to the confusion. As we saw in the examples, many of the interpretations of focus seem to have prosodic realizations in English. Even though many of these notions of focus may be independent of each other, in English the similarity of realization tends to obscure the potential for differences in interpretation.

Prosodic marking of focus, which seems so pervasive for English, is not universal. Work has been done on languages that are quite different from English with respect to how focus is expressed. Hungarian, for example, has been of great interest to linguists for expressing syntactically the same types of phenomena that English expresses prosodically. However, researchers on Hungarian and English take somewhat different phenomena as being of interest. The perspectives of researchers working on focus in Hungarian are not identical
to those of researchers working on focus in English. Additionally, no one has had a body of data that would provide the best comparison, that is utterances in both languages for the same contexts. Consequently, it is really still an open question as to whether what is being discussed as focus in English is the same as what is being discussed as focus in Hungarian.

The goal of this study is to resolve some of this confusion about focus by testing the idea that a number of distinct notions of focus may be necessary, each independent and each associated with a relatively simple linguistic realization. The complexity observed is, under this type of analysis, a result of the interaction of these simple components. Once this 'many types of focus' view is adopted, it becomes clear that in order to investigate one type of focus one must control for the effects of the others. That is the core of this study: investigating which interpretation is associated with which realization once the effects of other types of focus have been factored out. The inclusion of both English and Hungarian in the study is also intended to aid in the task of sorting out various types of focus by providing an opportunity for comparison between these two languages in which focus realization is quite dissimilar.

The first distinction that needs to be made toward this goal is to separate the prosodic, syntactic, morphological and lexical phenomena, which are part of the realization of focus, from notions like informativeness, givenness--newness and contrast that are part of the interpretation. With this crucial division between interpretation and realization clearly in mind, the tasks of this study are as follows:

1. Identify a set of interpretations of focus that are likely to be distinct and that can be productively investigated.
2. Identify a set of realizations likely to be associated with the interpretations identified in 1.

3. Develop a technique for eliciting relatively spontaneous speech that also has the following properties:
   a. Contextual factors related to interpretations can be manipulated in order to elicit tokens in which interpretations that often naturally co-occur are separated. This manipulation of the context should also enable more accurate coding of interpretations.
   b. Stimuli are visual so that the same stimuli can be used with speakers of different languages to collect maximally similar multilingual data without introducing translation artifacts.

4. Investigate which interpretations (from 1) are associated with which realizations (from 2) using the corpus (from 3) in which interpretations can be reliably coded and their effects on realizations separated due to the design of the corpus.

5. Compare results for English and Hungarian.

The following overview of the dissertation refers back to the list of tasks for the study just discussed and provides pointers to the sections of the dissertation which address these tasks. The interpretations and realizations in items 1 and 2 from the list of tasks above will be derived from consideration of the previous research. Three types of focus interpretations are particularly pervasive in the literature: information structure, new vs. old or given, and contrast. These three will be further investigated in this study. The focus realizations that seem from the literature to be connected to the most prevalent focus interpre-
tations are prosodic prominence and syntactic position, prosodic prominence being especially relevant in English and syntactic position being important in Hungarian. A discussion of interpretations and realizations of focus from previous work appears in Chapters 2 and 3.

Another major part of the study, the Animation Corpus, was developed to fulfill the requirements of item 3 above in providing a control of discourse factors through manipulation of the task context and in providing a means for collecting maximally similar spontaneous speech from different languages. Subjects were taped while performing interactive tasks that involved exchanging information about simple computer animations. The animations used geometric shapes and icons of common objects as characters. Each character had one of six colors that can be named by a basic color term in any language that has at least six basic color terms. The use of common objects and basic colors for characters in the animations reduced the amount of disfluency since subjects were unlikely to have difficulty in naming either the shapes or the colors. The combinations of shape and color provided enough unique characters that they did not have to be reused. Consequently “new” characters could always be completely new. Three different tasks were used in collecting the Animation Corpus to elicit types of speech ranging from short question-answer exchanges to more elaborate descriptive narratives. The corpus collection also took advantage of the relationship between wh-questions and focus to elicit utterances with particular foci. Details of the Animation Corpus are discussed in Chapter 4 and some representative transcripts appear in Appendix A.
As previously mentioned, prosody is an important component of the realization of focus, particularly in English. It is investigated in this study by examining a number of acoustic phonetic properties. Particular phonetic properties were investigated because of their association with significant prosodic phonological events. For example, greater duration, greater amplitude and high fundamental frequency (F0) are phonetic phenomena that all tend to be associated with the phonological component of focus, pitch accent. Measuring duration, amplitude or F0 is not the same as identifying phonological components such as pitch accents, prosodic phrasing or boundary tones because the mapping between the phonological level and phonetic implementation for prosody has not been completely worked out. However, the phonetic measurements provide indirect evidence for the phonological components, particularly when considered together because they measure phonetic properties generally associated with phonological components. In addition to considering duration and amplitude, this study also calculates the integral of amplitude over time. This measurement essentially combines duration and amplitude and gives a sense of overall energy across a unit. Duration and amplitude for English are discussed in Section 5.1.

For F0, I measure maximum F0, mean F0, the F0 at the point of maximum amplitude and F0 range. These measures are likely to be associated with pitch accents. I also measure peak offset which relates to the alignment of a peak relative to the point of maximum amplitude and should detect differences in pitch accent type. F0 for English is discussed in Section 5.2.

While prosody is the major mode of expressing focus in English, syntax is considered the dominant method of expressing focus in Hungarian. Hungarian is a discourse configura-
tional language which means that syntactic position is sensitive to discourse notions such as focus and topic rather than to grammatical notions such as subject and object. Consequently, most of the discussion of focus in Hungarian has been about the syntactic behavior and the interpretation of the Preverbal Focus Position. In Chapter 6, I examine three recent proposals about syntax and focus in Hungarian and consider how well they can account for the syntactic positions of occurrences focus in the Hungarian Animation Corpus Data.

There are four major results from this work. The first is demonstrating that at least one type of focus in English, the type I will call information structure focus, is marked on all the words within a focus constituent. This differs from a very common view that marking one word within a constituent serves to mark the whole constituent as focus. Finding marking for focus on all words in a constituent also calls into question the related idea that there is ambiguity about how large a constituent is marked as focus by pitch accent placement on a particular word. This study demonstrates more convincingly than previous work that there is marking on the whole focus constituent.

The second major result is showing that current syntactic accounts of focus in Hungarian do not cover all the patterns found in the Animation Corpus. The closest account, that of É. Kiss (1996), still fails to account for certain restrictions on the location of information structure focus. It appears that there are additional factors to consider, and I suggest what some of these might be. In particular, I suggest that other types of focus phenomena and metrical structure need to be considered in accounting for the distribution of information structure focus.
Another contribution of this study is the development of techniques used in collecting the Animation Corpus. This technique, using animations and drawings in a task, succeeded in eliciting a corpus that had useful control on discourse factors and vocabulary but remained relatively spontaneous. The manipulation of the task yielded a variety of desired types of interactions from simple question-answer exchanges to more elaborate descriptive narratives. The use of the animations and drawings as stimuli allowed for a technique that works crosslinguistically. The same tasks and stimuli worked equally well with English speakers and Hungarian speakers and would be easy to extend to other languages. The resulting corpora are as comparable across languages as is possible without using reading tasks and giving up the spontaneity of the speech. If one wants to examine the same interpretation of focus in Hungarian and English, this corpus provides a chance to do that by eliciting utterances in the two languages in the same context. This possibility did not exist for previous researchers who did not have such a corpus, so this study provides new information for comparing Hungarian and English focus.

A fourth contribution of the study is in the use of matched pairs for analysis. Matched pairs are pairs of items, in this study words, that share the same values for all relevant properties except the one being investigated. Particularly in investigating prosodic features, the matched pair analysis provides a way to make interesting comparisons without having to adopt a particular phonological approach. My use of this pairing facilitates the isolation of factors that could affect prosody and allows me to demonstrate that newness is not realized prosodically in English.
Chapter 2

Previous Work on Focus

‘That's a great deal to make one word mean,’ Alice said in a thoughtful tone.

‘When I make a word do a lot of work like that,’ said Humpty Dumpty, ‘I always pay it extra.’ (Carroll 1941; Carroll 1993)

The literature on focus is vast and confusing. This chapter will discuss major themes in the literature on focus, areas of agreement and points of controversy. This review is not meant to provide a comprehensive history of research on focus but rather a useful overview of key concepts. The chapter will start with a brief discussion of the most typical, most cited type of focus phenomenon -- the answers to wh-questions. The discussion of answers to wh-questions is intended to give the reader a feel for the type of data which the approaches surveyed in this chapter are interested in accounting for. In section 2.1, I will discuss some phonological/phonetic approaches to focus. In section 2.2, the discussion will cover two major conceptual types of accounts for focus: Movement, and Partition. The discussion of Focus Movement type accounts and Partition type accounts takes the interpretations of focus posited by these the various approaches as its primary subject matter. However, connections to phonological and syntactic realizations will be made as needed.
Answers to wh-questions are important in research on focus because these answers occur in one of the few contexts in which identification of focus is relatively uncontroversial. That there is a relationship between the focus of the answer and the wh-item of the question it answers is one of very few areas of wide-spread agreement in the research on focus. Specifically, the relationship between the two is that the focus of the answer must be the constituent that provides instantiation of the variable represented by the wh-item of the question. This relationship between wh-questions and their answers has been exploited by many (Gussenhoven 1984; Jackendoff 1972; Jacobs 1991b; Krifka 1991; Rooth 1985; Selkirk 1984; Vallduví 1990) as a diagnostic for the focus portion of the focus/ground partition. Examples (7)-(11) demonstrate this relationship between wh-questions and their answers and also show that the scope of the focus varies with the type of wh-item. Focus is represented in examples (7)-(11) and in much of the literature by CAPITAL letters. In (7) what represents a sentence so the focus of the appropriate answer is the entire sentential answer. In (8) what in combination with the pro-VP done corresponds to a VP focus in the answer. The answer in (9) focuses everything but the direct object Tommy, and so on.

(7) What's happened?

PAPA HAS GIVEN TOMMY A GUN. (Gussenhoven 1984)

(8) What's Papa done?

Papa HAS GIVEN TOMMY A GUN. (Gussenhoven 1984)

(9) What happened to Tommy?

PAPA HAS GIVEN Tommy A GUN. (Gussenhoven 1984)
(10) What's Papa done to Tommy?

Papa HAS GIVEN Tommy A GUN. (Gussenhoven 1984)

(11) What's Papa given Tommy?

Papa has given Tommy A GUN. (Gussenhoven 1984)

It has also been noted that the minimal answer to a wh-question is the focus alone as in (8). That is, the focus is a required part of the response. This can be seen by comparing the answer in (12) with the longer but equally acceptable answer in (13).

(12)a. What doesn’t the boss like?

b. BROCCOLI (Vallduví 1991)

(13) The boss doesn’t like BROCCOLI.

The focus of answers to wh-questions have been widely argued or assumed to be marked by prosodic prominence in English and by word order in discourse configurational languages such as Hungarian. These two types of realization will be discussed in the next section.

2.1 Approaches to the Realization of Focus: Prosody and Syntax

There are many possible realizations of focus, but the two most discussed are phonological, particularly prosodic and syntactic. This section will discuss some approaches to accounting for the prosodic phenomena associated with focus, mainly in English, and will
discuss some approaches to accounting for syntactic phenomena associated with focus, mainly in discourse configurational languages. While this section will be primarily concerned with what happens in the syntax or the phonology once a constituent is identified as focus, it is impossible to entirely divorce discussion of realization from discussion of interpretation. Especially if one imagines, as is advocated in this study, that various interpretations may have different realizations. Consequently, brief mention of relevant focus interpretations will be made as necessary in this section but a full discussion of them will appear in section 2.2.

### 2.1.1 Prosody

Most researchers agree that the prosody of sentences reflects the speakers’ intended focus, but how the focus is represented is an area of considerable controversy. I will call the two major approaches to this question, following Ladd (1996), the ‘structure-based’ approach and the ‘highlighting-based’ approach. The highlighting-based approach could be characterized as a one to one mapping of accent to focus. Under the highlighting-based approach every focus has an accent and every accent marks a focus of some sort. Consider as an example of this approach Bolinger (e.g. 1958; 1972; 1986; 1989), an early and persistent proponent of the highlighting approach. Bolinger (1986; 1989) argues that words are accented because they are of interest. According to him, words can be of interest for a variety of reasons, including some of the notions mentioned above. The notion of highlighting for interest is not necessarily linguistic. Bolinger considers interest a more general cognitive property.

(14) Raw fish is good for you but after all, who likes raw FISH? (Bolinger 1986)
(15) Raw fish is good for you but after all, who likes THAT? (Bolinger 1986)

Bolinger argues for these example that the accented “fish” and “that” are neither new, nor informative nor evoking contrast. Rather “fish” and “that” are serving a pointing function with the accent and are of interest because of this pointing function. Notice that the relevant unit for the highlighting-based approach is the word and not larger constituents.

The structure-based approach (Gussenhoven 1984; Ladd 1980; Schmerling 1976, inter alia) differentiates between the distribution of accents and the distribution of focus. In the structure-based approach the narrow focus case is the same as for the highlighting-based approach: if a word is the focus it receives an accent. The term ‘broad focus’ (as suggested by Ladd 1980) represents the idea that a constituent larger than a word can be focussed and it is on this case that the two approaches diverge. Broad focus requires an account of how accent placement on a word can represent a focus on a constituent that is bigger than a word. Accent placement for broad focus is a central concern in the accounts of Ladd (1980; 1983), Gussenhoven (1984) and Selkirk (1984). All of these studies propose some form of focus ‘percolation’ in which various factors such as syntactic structure, phonological phrasing and predicate-argument -modifier status affect how far the focus marking effect of an pitch accent can extend beyond the word that it is aligned with. For example, one of Gussenhoven’s domain formation rules says that a focus predicate followed by a focussed argument can form a single accenting domain. So in a sentence such as “Joe bought a DOG”, the single pitch accent represented by the capital letters on dog can be used to mark as focussed the entire VP “bought a dog”.
The structure-based approach is the type of focus phonology assumed by many works concerned primarily with the semantics of focus and focus particles (e.g. Krifka 1991; Rooth 1985). Steedman (1997), on the other hand, takes an approach that can be seen as a combination of the structure-based approach and the highlighting-based approach. He uses intonational boundary tones as the delimiters of the focus domain (in his terminology the ‘rheme’) and a single accent within that domain marks the whole domain as focussed. The placement of that accent within the domain (rheme) marks a particular word as ‘contrast’ which, if one takes contrast to be a type of focus, is a perfect example of the highlighting-based approach.

2.1.2 Syntax

Various syntactic constructions in English have been claimed to mark focus (e.g. Rochemont 1989). The list of such constructions includes extraposed relatives, inversion, presentational there sentences and it-cLEFTs. The it-cLEFT in particular has been taken to be a canonical example of marking focus syntactically. While there is no doubt that the fairly exotic syntax of English it-cLEFTs is being deployed to mark something, it is not clear that what it is marked by it-cLEFTs is the same as what is marked by prosodic prominence in English, even though both have been called focus. Most it-cLEFTs that are given as examples of focus are what Prince (1978) calls “Stressed-Focus” it-cLEFTs. Stressed-Focus it-cLEFTs such as (16) not only have the it-cLEFT syntax, but the ones cited as examples of focus are also typically assumed to have a substantial pitch accent on the clefted element (= BETH in (16)). Since it is widely accepted that pitch accents mark focus in English, one might wonder how it would be possible to conclude that it is the syntax of it-cLEFTs doing
the focus marking rather than the prosody. Perhaps if it-clefts unfailingly had the major
accent of the sentence on the cleft, arriving at this conclusion would be understandable,
since the putative focus marking and the prosodic focus marking would be inseparable.
However, this is not the case. Examples (17)-(19) show that the pitch accent can shift in
it-clefts just as it can in more mundane declaratives, and changes in focus in it-clefts
accompany changes in pitch accent location in just the same way as in other constructions.

(16) It was BETH who gave her dad a dog.

(17) It was Beth who gave her DAD a dog.

(18) It was Beth who gave her dad a DOG

(19) It was Beth who GAVE her dad a dog.

Also, it has been observed by Ball and Prince (1977) Prince (1978) and Delin (1989) that
many instances of it-clefts are not the “stressed-focus” type. They are what Prince calls
“informative presupposition” it-clefts and the main sentence accent in this type of it-cleft
is located in the clause (e.g. (20))

(20) It was for a career that he no longer wanted that Alfred had made all these SAC-
RIFICES

Other alleged focus constructions in English have been argued to mark discourse familiar-
ity. That is, how new or old the item is to the discourse. For example, Birner (1992)
argues that inversions such as in her example 1a reproduced here as (21), mark the pre-
posed element as older in the discourse than the postposed element.

(21) <also eliminated> is <the expense of buying costly chemicals> [WOODEXTRA,
What we see with the it-clefts and with inversion is that the sort of focus being marked is not the same as the type of focus being marked prosodically in answers to wh-questions. This supports the perspective of this study, that more than one type of focus is necessary to account for the range of phenomena that have fallen under that label of focus.

While syntactic marking in English is perhaps something of a peripheral phenomenon, syntactic marking of focus is a central part of focus-prominent languages such as Hungarian (Brody 1990; É. Kiss 1995c; Horvath 1986; Horvath 1995), Basque (Uriagereka 1988), Bulgarian (Rudin 1990), Korean (Choe 1995), and most Chadic languages (Horvath 1995). These languages have a structural focus position, often adjacent to the verb (at least at some point in the derivation). In recent analyses, many of these languages are taken as having a functional projection headed by focus and movement driven by the need for licensing or assignment of a +focus feature provided by the verb.

Finally, consider a prosody/syntax combination. Vallduví (1990) argues that for Catalan, rather than having a structural focus position, syntactic movement shifts material out of the clause, dislocating it both right and left and leaving the focus item in clause final position where the main sentence accent is obligatorily located. One could view this as either marking the focus by syntactic means or as using the syntax to arrive at the appropriate prosodic marking.
2.2 Approaches to the Interpretation of Focus

2.2.1 Focus Movement

One approach to focus is to treat it as movement, similar or identical to wh-movement. The Focus Movement approach takes focus to be some type of operator that moved either at surface structure or at LF to attain appropriate scope. Some inadequacies of the Focus Movement approach for languages like English that do not have overt focus movement were demonstrated by Rooth (1985) with examples such as his (22) and (23).

(22) They only investigated the question whether you know the woman who chaired THE ZONING BOARD. =(Rooth 1985, example 57)

(23) *Which board did they investigate the question whether you know the woman who chaired ei? =(Rooth 1985, example 58)

The focus movement that would have to be posited at LF for THE ZONING BOARD in (22) is not possible as wh-movement for which board in (23) at LF or at S-structure levels. These examples show that a movement analysis for focus would require it to have a type of movement unlike any other in that it does not obey island constraints. Rather than propose a new type of movement that can violate island constraints, relational semantic approaches that can interpret focus in place (Rooth 1985; von Stechow 1981; von Stechow 1991, inter alia) and discourse/pragmatic accounts that process focus on a separate level from LF (Vallduví 1990; 1995) have gained favor in accounting for languages such as English that do not have structural focus positions. These relational semantic and discourse/pragmatic accounts do not require as input an LF form with the focus item(s) in
a moved position, so much of the motivation for LF focus movement in English has evaporated.

However, for discourse configurational languages with structural focus positions, the movement analysis is quite intuitive. In these focus-prominent languages such as Hungarian, focus movement does obey subjacency and looks very similar to wh-movement. É. Kiss (1987) in discussing focus-prominent languages suggests that there are at least two types of focus:

- it [focus] can denote the sentence part carrying new information, and it can also mean an operation expressing identification.
- it is the focus operator that is associated with a particular structural position in most discourse configurational languages, as well.

In addition, in these languages, the item in the structural focus position is in a similar position to other items that are uncontroversially taken to be operators (e.g. quantifiers), so the interpretation of focus as an operator seems more plausible.

### 2.2.2 Partitioning

Many semantic and discourse accounts share a view focus as a part of a partition. This is in fact the current dominant view among those doing research in this area. Accounts differ on the details of the partition, what the function of the partition is, how the pieces of the partition relate to each other and what properties are attributed to the various pieces of the partition. The two major modes of partitioning are the “topic/comment” or “theme/rheme” split (Chomsky 1977; Gundel 1974; Gundel 1985; Halliday 1967; Hockett 1958; Lambrecht 1994; Prevost 1995; Prevost and Steedman 1994; Steedman 1997; Strawson 1964, inter alia) and the “focus/ground” or “focus/presupposition” split (Akmajian 1973;
Chomsky 1972; Jackendoff 1972; Jacobs 1991a; Prince 1981a; Prince 1984; Prince 1986; Rochemont 1989; Rooth 1985; Vallduví 1990; von Stechow 1981, inter alia). There are other terms used for the components of these partitions, and there are other studies than the ones cited that use similar divisions. To escape the clumsiness of multiple terms, I will generally refer to the “topic/comment” or “theme/rheme” split simply as the topic/comment split and to the “focus/ground” or “focus/presupposition” split as the focus/ground split.

2.2.2.1 Topic/Comment

In the topic/comment split, the topic is generally agreed to be what the sentence is “about”. Unfortunately, “aboutness” is difficult to identify, define or test for. Tests that have been proposed (Gundel 1974; Reinhart 1982), for example the ‘as-for’ test, the ‘said-about’ test and the ‘what-about’ test, have proven unreliable in identifying even all cases of syntactically marked topicalization in English (Gundel 1974; Prince 1984; Vallduví 1990; Ward 1985). Another suggestion about identifying topics made by Halliday (1967) is that in addition to “aboutness”, topics are the first element in a sentence. This is also problematic and both misses cases of non-initial topics and attributes topichood to elements such as fronted wh-words that do not seem plausible candidates for what the sentence is about.

Steedman (1997) adds to the topic and comment, which he calls ‘theme’ and ‘rheme’, “a distinction between the contrastive part(s) of either information unit, and the rest.” He calls the portion of each unit associated with a pitch accent the ‘focus’ and the portion not associated with a pitch accent or boundary tone(s) the ‘background’. Note that while the
rheme always has a pitch accent/focus in this system the theme is allowed the option of going accentless. Although Steedman uses the topic/comment partition the semantics he proposes is based on the structured propositions approach (Creswell 1973; Creswell 1985; Jacobs 1991a; von Stechow 1981; von Stechow 1991, inter alia) and the alternative semantics of Rooth (1985) which use the focus/ground partition. Steedman’s theme/rheme partition differs from the focus/ground partition mainly in those cases in which there is additional material in the rheme that would not qualify as part of the focus in a focus/ground split. Consequently, in most cases his approach is rendered compatible with work in alternative semantics and structured propositions by treating rheme and focus as equivalent and theme and ground as equivalent. Steedman interprets the theme as referring to the alternatives set associated with the utterance. Steedman’s alternatives set for a sentence is constructed by λ-abstraction over a version of the proposition represented by the sentence which has had the rheme replaced by a variable. This in essence creates a function from the logical representation of the sentence which can take as an argument the rheme or other items like the rheme that could be alternatives to it. To form the alternative set these “potential” rhemes are plugged into the function built with the lambda abstraction operator to produce a set of propositions that are ‘alternatives’ to the proposition represented by the sentence. One could also substitute the existential quantifier ∃ for the λ-abstraction operator and take as the alternatives set referred to by the theme, a set of propositions that could instantiate this existentially quantified proposition. The rheme is interpreted as a restrictor on the alternatives set, restricting it to the alternative in which the λ-abstraction or the existentially quantified proposition is instantiated by the actual rheme of
the sentence. Examples (24) and (25) from Steedman (1997) illustrate these components of his analysis.

(24) Q: Well, what about MUSICALS? Who likes them?

A: (MARY) (likes musicals).

H*L  L+H* LH%

=(Steedman 1997, example 3]

(25) Q: Well, what about MARY? What does she like?

A: (Mary likes) (MUSICALS)

L+H* LH%  H* LL%

=(Steedman 1997, example 4]

These examples show two different theme/rheme articulations for the same string. Example (24) shows the rheme “MARY” is marked by the H*L intonation pattern while “MUSICAL” is marked as rheme by the H*L pattern in (25). The themes are marked by the L+H*L H% intonation patterns. The existentially quantified proposition corresponding to (24) appears in (26).

(26) $\exists x. \Diamond \text{like}' x \text{mary}$

(27) shows a possible alternatives list for the same sentence.

(27) $\{ \Diamond \text{like}' \text{jazz'mary}', \Diamond \text{like}' \text{opera'mary}', \Diamond \text{like}' \text{musicals'mary}' \}$

The placement of the accent within the theme and within the rheme is interpreted by Steedman as marking contrast: “It is important to note that it is all and only the material marked by the pitch accent(s) that is contrasted.”
In discourse configurational languages that are topic prominent, the topic has been argued to be the “notional subject”. In other words, in such languages the main predication of sentences is typically a topic-predicate relation rather than the (grammatical) subject-predicate relation found in English. In these languages, the topic has many of the syntactic properties associated with the grammatical subject in English and would presumably be easier to identify, since there would be some potential for developing robust syntactic tests of topichood.

The topic/comment split will not be as crucial to this study as the focus/ground split because the topic/comment split alone does not isolate the items normally discussed as “focus.” These items are simply part of the comment. However, Steedman’s approach was presented in some detail because it makes additional distinctions for contrastive vs. non-contrastive elements, and because his treatment is so similar to certain approaches that use the focus/ground partition.

Having seen some variants of the topic/comment partition, we have also seen some indication that there are a number of proposed properties and interpretations for the components of the partition. Among these are:

1. The topic is what the sentence is “about.” (Gundel 1988; Halliday 1967; Hockett 1958; Lambrecht 1994; Reinhart 1982; Vallduví 1990)
2. The topic is an index to the hearer’s knowledge store, the direction for where to store the proposition. (Reinhart 1982; Vallduví 1990) (see also discussion of Vallduví (1990) in section 2.2.2.2)
3. The Topic is the Subject of the main predication in sentence (in topic promi-
4. The topic (theme) refers to the alternatives set associated with the utterance.  

(Steedman 1997)

How is the topic/comment partition realized? One suggestion was that the topic is sentence initial (Halliday 1967; Vallduví 1990). While Halliday’s extreme position that all initial elements are topics clearly doesn’t work, it is fairly common for languages to have constructions that place topics in positions near the beginning of sentences. Topicalization in English is an example of such a construction. This might be taken as support for Vallduví’s more moderate position, which places topics in initial position when they occur. However his position may also be overstated. In some topic-prominent languages, there are left-dislocated constituents that can appear to the left of the structural topic position, and which may or may not be interpreted as topics. Another potential realization of topics is deaccenting, which seems to occur in some languages (É. Kiss 1995a). While in other languages topics are thought to receive different types of pitch accents or tunes than non-topics and/or to be separated from the rest of the utterance by a prosodic boundary.

2.2.2.2 Focus/Ground

In the focus/ground split, Focus is one piece of the partition and the other part of the partition is what Prince (1981b; 1986) terms an open proposition. The open proposition resembles the original utterance but has a variable in place of the focus.¹

---

¹ In other accounts, the component which is analogous to the open proposition has variously been termed presupposition (Akmajian 1979, inter alia; Chomsky 1972; Jackendoff 1972), topic (1991; Sgall, Hajicová, and Panevová 1986, inter alia), presupposition skeleton (Rooth 1985), and focus frame (Partee 1991). The structured propositions adopted by Jacobs (Jacobs 1991a) and Krifka (Krifka 1991) support a similar partition.
For example, in (59) from Prince (1986), (59b) represents the open proposition of (59a) with \textit{SHIRT} as the focus.

\textbf{2.2.2.2.1 Interpreting the Focus/Ground Partition: Information/Informativeness}

Given the focus/ground partition, one answer to the question of how the parts are interpreted is that the focus is the “informative” part of an utterance (Danes 1968; Dretske 1981; Firbas 1964; Firbas 1971; Vallduví 1990). Vallduví (1990) relates this notion of informativeness to reducing uncertainty in the hearer's knowledge store. The information of a sentence is the part of the propositional content that makes a contribution of knowledge to the hearer's knowledge store. Unlike propositional content the informativeness of an utterance or a part of an utterance is dependent on the state of the hearer's knowledge.

Understanding utterances is a process of incrementally updating one’s knowledge store. The updating procedure could be different for the same propositional content in situations where the hearer's knowledge store differed. Vallduví’s model of the knowledge store is based on Heim’s file card semantics (Heim 1983).

Example (60) from Schmerling (1976) was spoken in a context in which it was known that Truman was ill.

(60) Truman's DIED. (Schmerling 1976)

\[ \Lambda x, x=\text{Truman}[\Phi[x \text{ died}]] \]
Under Vallduví’s analysis, when (60) is uttered an entry already exists on a file card in the hearer’s knowledge store indexed on Truman. The notation “Ax,x=Truman” represents that Truman is a type of non-focal material Vallduví calls LINK. The focus-ground structure of this example instructs the hearer to enter the focus "$\Phi[x\ died]" to his/her knowledge store by entering it on this preexisting file card.

(61) JOHNSON's died. (Schmerling 1976)

$\Phi[Johnson's\ died]$

This example is analyzed by Vallduví as an all focus sentence. According to Schmerling (61) was uttered with no prior discussion of Johnson and no knowledge that he had been ill. The hearer is instructed to enter the whole proposition into the his/her knowledge store on a new card. There is no preexisting file card since there was no prior discussion.

2.2.2.2 Interpreting The Focus/Ground Partition: Given/New

There have been almost as many definitions of given and new as there have been of focus. One is a simple dichotomy between given and new, where new is defined as “what the speaker assumes he is introducing into the addressee’s consciousness by what he says” (Chafe 1976), or as not having an antecedent in the discourse (=not c-construable) (Rochemont 1989), or by more or less similar notions of familiarity predictability or shared knowledge. When the term focus is used based on a new/old distinction, focus is generally taken to be new and ground is taken to be old. This simple dichotomous version of new/old is also the notion adopted by Selkirk (1984; 1995). For Selkirk, some part of the focus must be new, that is, new to the discourse, although some material within the focus.
can be given. A problem with using this type of new/old distinction for the answers to wh-questions is that the constituent that answers to the wh-question can be familiar to the hearer and salient in the discourse. It has been argued, for example by Vallduví (1990), that what is “new” in the answers to wh-questions is that a particular focus fills a particular open proposition or ground. In other words that the newness is a relational matter between the pieces of the partition rather than being based on the referential status of one piece of the partition. This relational notion of newness is what Vallduví and others call information or informativeness.

Two questions arise with respect to determining new/old status, how familiar or unfamiliar does something have to be to be old or new and what is it new or old with respect to? On the first of these questions Prince (1981a) argues that a binary distinction is inadequate to account for a number of discourse phenomena and proposes the ranking of assumed (hearer) familiarity shown in (28). In this ranking, Evoked material has explicitly appeared in the previous discourse or is situationally available and salient, Unused material is presumed to be known to the hearer but not evoked in the current discourse, Inferables are presumed to be retrievable by the hearer from salient shared knowledge, Containing inferables are a sub-case of inferables in which the inference is supported by some part of the inferables item’s description (for example “one of the top candidates” is inferred from “the top candidates”), Brand-New Anchored entities are Brand-New but connected to some salient entity in the discourse (for example the speaker in “someone I’d never met”), and finally Brand-New items are completely novel to the discourse.

(28) Evoked>Unused>Inferrable>Containing Inferable>Brand-New
In answer to the question of what something might be old or new with respect to, Prince (1992) proposes a distinction between discourse familiarity and hearer familiarity. Of the three useful combinations: discourse old/hearer old, discourse new/hearer new and discourse new/hearer old. Of these, the first two conveniently correspond to the two endpoints of Prince’s 1981 ranking, Evoked and Brand New.

2.2.2.2.3 Interpreting the Focus/Ground Partition: Contrast

Contrast has sometimes been treated as separate from other notions of focus, for example by Vallduví (1990), Steedman (1997), Prevost and Steedman (1994) and Prevost (1995). Other researchers have seen as a contrast as a optional property of focus, for example Kenesei (1986; 1994). Yet others take focus as being necessarily contrastive (e.g. Dretske 1972; Rooth 1985; Szabolcsi 1981; Szabolcsi 1983). Regardless of the view of contrast, the most consistent underlying notion is that of exclusion or exhaustiveness.

A point about contrast, observed by Bolinger (1986; 1989), Schmerling (1976), Kenesei (1986; 1994), Vallduví (1990), and many others, is that contrast seems stronger when the salient alternatives are limited and weak or absent when there are many possibilities. One strong contrast case is in the answers to questions that explicitly state two alternatives as in (29) from Prevost (1995).

---

2. Actually, while Steedman and Prevost conceptually separate contrast from new/old and information structure, their system requires that all rhemes (=foci) contain a contrastive element. Steedman in particular also uses the term “new” for the items which have pitch accents, although clearly these items need not be discourse new.
(29) Q: Would you prefer the BROWN ale or the PALE ale?
   A: I’d like the PALE ale. = (Prevost, 1995, example 26),

In this example, the question makes the salient possibilities be “pale” and “brown”. While
the respondent could answer with some other choice, there is a strong expectation that he
or she will answer with one of these choices. With the limited choices the feeling of con-
trastiveness is strong. Compare that to (30) in which one can imagine an infinite number
of things that could be wrong with Agnes, and from which one does not get a sense of con-
trast in the answer.

(30) Q: What’s wrong with Agnes?
   A: She has a MIGRANE.

Kenesei argues based on this sort of evidence that contrast is available when a focus oper-
ator quantifies over a closed set of individuals. In this case, identification of the individual
for which the proposition holds also picks out a closed complement set of individuals for
which the proposition does not hold. So in , since there are only two contextually relevant
alternatives, a complement set of one {brown ale} is created by selecting pale ale. In (30),
however, with essentially unlimited possibilities, no sensible complement set can be
formed as a by-product of selecting “migraine” so it is not contrastive but merely ‘identifi-
cational.’

Prevost (1995) and Prevost and Steedman (1994) also take exclusion as a crucial part of
contrast. Words are marked as contrastive in their system if the property represented by
that word excludes possible referents. In (29), the property represented by the word ‘pale’
excludes ‘brown ale’ as a possible referent. In comparison, since both choices are types of ale, the word ‘ale’ would not qualify as contrast.

Contrast has been considered a derived notion in many approaches to focus (e.g. Vallduví (1990), Prince (1984)). However an argument against abandoning contrast as a basic notion is that there are a few phenomena that seem to be strictly contrastive and not particularly related to other notions. Consider the example in (31). When pronouns are accented, their referent is switched and is not the antecedent that would otherwise have been expected. One approach to what is “expected” as antecedents to pronouns is Centering Theory (Grosz and Sidner 1986). Without going into the details of Centering, given an appropriately ordered list of potential antecedents, Centering Theory could predict that the pronouns *he* and *him* in (31) and (32) have the antecedents indicated by the indexing in (32) rather than those in the joke in (31). The pronoun resolution in (31) illustrates the “anti-default” effect of accenting pronouns.

(31) Johnj called Billi a Republican and then HEj insulted HIMj.

If the pronouns are deaccented John is doing the insulting in the second clause as in (32).

(32) Johnj called Billi a Republican and then hej insulted himi.

Ladd suggests a special reciprocal interpretation for such accented pronoun interpretations.

(33) JOHN likes MARY and BILL SUE.
The verb ellipsis case in (33) does not involve anything reciprocal but the structural parallelism seems to similarly facilitate particular contrastive interpretations. List of numbers or colors seem to work similarly in referring for contrast to prior utterance patterns as in (34).

(34) green orange blue red, green YELLOW blue red...

2.2.2.2.4 Interpreting the Focus/Ground Partition: Alternatives

Another type of answer to the question of how the focus/ground articulation is interpreted is that focus evokes a set of alternatives, which is carried in the semantics as a second semantic value. In Rooth (1985), this is called the focus semantic value. This discussion should be fairly familiar from the discussion of Steedman (1997) in section 2.2.2.1. The focus semantic value of a sentence is derived from the ordinary semantic value by making a substitution in the position corresponding to the focused phrase.

(62) MARY likes Sue.

ordinary semantic value = like(m,s)
focus semantic value = \{like(x, s) \mid x \in E\}, where E is the domain of individuals.

These alternatives from the focus semantic value are then available as the domain for focus sensitive operators like “only” and “even”. The Structured Propositions approach (e.g. Jacobs 1991, Krifka 1991) uses a different formal mechanism, but the underlying intuition is very similar. These approaches are more directed towards using the focus/ground partition to provide scopal domains for the interpretation of various focus sensitive operators than toward interpreting focus and ground elements themselves. Generally,
when there is no overt focus operator, these approaches posit an invisible one, such as a speech act level ‘ASSERT’ operator or something like an invisible ‘only’. An additional interesting point about these approaches is that the consideration of alternatives at least implies rejection of all of the alternatives except the focus. So the alternatives-type analysis would seem to at least suggest a property of exclusion or exhaustiveness for focus which fits more naturally with a contrast based notion of focus than with an information or new/old based one.

2.3 Summary

In reviewing the previous work there seem to be three very pervasive interpretations of focus: information structure, new/old, and contrast. Most other approaches/concepts/properties discussed in this section could be viewed as related to one of these three. For example alternatives and exclusion could both be viewed as part of contrast. In fact, “exclusion of relevant alternatives” is a fairly reasonable definition of contrast. The notion of focus as an operator, involves either exclusion, which could be taken to be contrast, or identification, which seems very similar to informativeness. Given the prevalence of information structure, new/old and contrast as interpretations of focus, these will be the notions investigated in the study. Discussion of the issues involved in identifying and defining these notions for the study are discussed in Chapter 3.

In terms of realization, variations on the scheme focus↔accented and non-focus↔deaccented are very common. Deaccenting has especially been argued for in relation to Old items. There is some controversy about whether the phonological account should support
a one-to-one mapping between accenting and focus interpretation or whether a ‘percolation’ account in which a single accent marks a larger domain is more reasonable. It has also been suggested that different pitch accent types and/or different intonational tunes mark focus and non-focus items. Accenting is clearly the obvious realization to investigate for English. Consequently, this study will look at a number of acoustic features that are indicative of different accenting and intonational patterns.

Syntactic realization is in the form of special syntactic positions for focus in many languages, e.g. Hungarian. In English, the use of syntax is more peripheral, prosody being generally recognized as the main way to mark focus. The English constructions that have been claimed to mark focus do not show a consistent syntactic position for it. Some of these constructions, for example it-clefts and inversion, seem mainly to be marking exclusion or new/old. Since syntactic focus marking is more peripheral in English, this study will concentrate on the type of syntactic marking found in Hungarian.
Chapter 4

The Animation Corpus

4.1 Introduction

The Animation Corpus was collected using simple animations and diagramatic drawings to engage speakers in contextually restricted question-answer and description dialogs. There were three basic tasks: the Question-Answer Task, the Directed Description Task and the Unrestricted Description Task. In these tasks, the animations and/or drawings were manipulated to bias what the speakers could consider DISCOURSE OLD vs. DISCOURSE NEW, what characters could be considered RELEVANTLY NON-UNIQUE, and what could be construed as INFORMATION STRUCTURE FOCUS or GROUND, as discussed in Chapter 3. In particular, the materials were designed to separate RELEVANT NON-UNIQUENESS, DISCOURSE NEW/OLD and INFORMATION STRUCTURE, since examples of these factors occurring on different items is more rare in natural discourse than having them co-occur on the same item. Since separating these factors is extremely useful in examining their individual contributions, it is profitable to set up contexts in which the frequency of occurrence of such utterances is greatly increased. As an example of how these factors can be separated, an animation might bias the agent of an action towards being new by not having it appear in any prior animation in the sequence. The recipient of the action could be RELEVANTLY NON-UNIQUE by having another object parallel its behaviour and/or share a property with it such as shape or color. Thus the agent character can be set up to be easily interpreted as DISCOURSE NEW but not RELEVANTLY NON-UNIQUE, while the recipient character can be set up to be easily interpreted as RELEVANTLY NON-UNIQUE and DISCOURSE OLD. In the Question-Answer Task and the Directed Description Task, the accompanying drawings are used to induce questions
about particular participants in the animation action, thereby shifting the INFORMATION STRUCTURE FOCUS of the response.

4.2 Characters

FIGURE 1. Selection of Animation Task Characters

4.2.1 Shapes

The characters used in the animations were taken from two international graphic symbol dictionaries (Dreyfuss 1984; Modley 1976). These simple icons have several attractive features. First, although not realistic they are highly recognizable. Informal testing of icons eliminated the few that caused difficulty in recognition. Since they are so recognizable task participants do not hesitate in naming them, and the possibility of additional discussion about the identity of objects is virtually eliminated. A second advantage of the icons is that since they were designed for and are used on signs in places like international airports, they have already been “tested” for cross linguistic effectiveness. The Hungarian speakers taped
for this study were residing in the Philadelphia Pennsylvania area at the time the corpus was collected, so there was little risk in assuming that these people would be familiar with the objects depicted as well as this type of iconic representation. With less well traveled individuals from more remote locations this assumption might be more problematic. A third advantage was that the icons were not complex drawings in the sense of having a lot of parts and were therefore relatively easy to manipulate in the animations. And a fourth very important advantage of the icons was that they did not constitute any kind of a natural or predictable set. In early versions of the task only geometric shapes were used and the speakers often seemed to treat these items as at least inferrable or possibly even DISCOURSE OLD, which caused problems for the manipulation of DISCOURSE NEWNESS/OLDNESS. The problem was especially serious since speakers did not even seem to be consistent about their treatment of the discourse status of the geometric characters. Rather the speakers treated the geometric characters as inferrable or not as they pleased and DISCOURSE OLD or DISCOURSE NEW as they pleased, leaving too much uncertainty about any given character’s discourse status. The icons used for characters were airplane, anchor, apple, ax, bed, bicycle, book, bottle, broom, bus, butterfly, camera, car, chair, (ear of) corn, cup, envelope, fork, (wine) glass, grapes, hammer, hand, hat, heliocopter, house, key, lighthouse, lock, motorcycle, package, paperclip, sailboat, scissors, sewing machine, shopping cart, shoe, suitcase, teapot, telephone, television, tooth, tree, truck, umbrella, and wrench. Some examples appear in Figure 2.

3. Even very isolated populations usually have some tradition of iconography that could be used as a source for suitable character.
FIGURE 2. Selection of Icons used as shapes for Animation Corpus characters
The geometric shapes circle, square, rectangle, oval and triangle were also used as characters. The diamond shape which was used in earlier versions of the animations was often confused with both the oval and the triangle and so was not used in the final version of the animations.

### 4.2.2 Colors

Six colors were used for the characters. The various colors provided a property that could be manipulated for **RELEVANT NON-UNIQUENESS**. This made six times as many characters available which helped ensure **DISCOURSE NEWNESS** and made the animations and drawings more attractive and interesting for participants in the tasks. In early version of the task, speakers had difficulty deciding on names for colors and distinguishing some of the colors from each other. To address the problem of the colors being difficult for the participants to name or distinguish, colors were adopted that fell within the focal range for the color terms white, black, yellow, red, green, and blue in English as discussed by Berlin and Kay (1969) in their seminal work on basic color terms. These focal ranges were shown by Berlin and Kay to be valid crosslinguistically so that any language with six basic color terms will name essentially the same colors. Choosing six color terms worked well for this corpus because both of the languages being investigated have at least six basic color terms. Choosing colors that correspond to these basic color terms completely eliminated difficulties with naming colors. Also, it is clear from Berlin and Kay’s work on basic color terms and subsequent psychological work that for a set of six colors, the ones that are named by the basic color terms will be the most distinct from each other and easiest to distinguish. And in fact, adopting the focal colors in the animations completely eliminated participants’ difficulties in distinguishing the colors. In addition, task participants found these focal colors extremely easy to remember. They showed no difficulty remembering colors of objects in the animations even on the few occasions when they did have difficulty remembering
shapes or actions.  

4.2.3 Actions

Each animation used one of three possible main types of action, which could be nicknamed eat, push and hit. In the “eat” action, the eater appears to develop a wedge shaped “mouth” as it nears the eatee. As the eater passes on top of the eatee, the eater’s “mouth” disappears, seeming to close, and the eatee disappears, seeming to have been eaten. In the “hit” action the hitter moves to the hittee and appears to rebound off of it. In the “push” action the pusher moves up to and touches the pushee then they move together for some distance across the screen while still touching. This gives the appearance of the pusher propelling the pushee. Another action, of appearing to pass through or under a brick “wall,” was used in some of the animations prior to the main eat, hit or push action(s) to cause some of the characters to be discussed prior to the main action(s). Each sequence of animations contained roughly equal numbers of eat, hit and push animations. Animations with more than one action always had either the “pass through wall” action followed by an eat, hit or push action, or a “pass through wall” action followed by two eat, hit or push actions. There was no mixing of eat, hit and push actions within an animation.

4. The only minor exception was in participants viewing the diagramatic drawings occasionally identifying the yellow objects as being orange. This problem was caused by a limitation in the capabilities of the color printer which sometimes printed yellow too dark. However, when yellow objects were identified as orange, task partners seemed to have no comprehension problem and the task proceeded smoothly. Since both English and Hungarian have basic color terms corresponding to “orange” this slight difficulty with color printing did not damage the functionality of the task.
4.3 Task 1: Question-Answer

In the Question-Answer Task one participant, the VIEWER, watches the animation while the other participant, the QUESTIONER, has a notebook of diagramatic drawings (static drawings) that correspond to actions in the animation. Examples of the QUESTIONER’s static drawings are shown in Figures 3 through 7.

FIGURE 3. Static drawing: “pass through the wall” color unknown
FIGURE 4. Static drawing: “eat” eatee unknown

FIGURE 5. Static drawing: “hit” hittee unknown
FIGURE 6. Static drawing: “hit” hitter unknown
In the QUESTIONER’s drawings, information is missing and is represented in the drawing by one or more question marks. In the majority of the drawings the question marks represents a missing character. For the drawings corresponding to the “pass through wall” action, the question mark represents either a missing character or a missing color (as in Figure 3). In the drawings that correspond to the animations designed to elicit focus particles, the question mark could also represent a portion of a character’s path. The QUESTIONER’s task is to find out the missing information represented by the question mark(s) in each drawing by asking the VIEWER. The QUESTIONER is not instructed to ask wh-questions, but in fact participants figure out very quickly that this is the most useful thing to do. An animation may have between one and three corresponding static drawings, depending on the number of actions in the animation. For example, an animation that has a red teapot pass under/through the brick wall, then a blue car hits a red cup, then it hits the red teapot,
will have three corresponding static drawings to elicit three questions, one for each of the actions. The first static drawing might be a drawing of a teapot with an arrow pointing toward a brick wall and a question mark in place of the color of the teapot (Figure 8). This usually elicits a question such as “what color was the teapot that went through the wall?” A second static drawing could show the red teapot, a red cup and a question mark with an angled arrow depicting the hit action on the red cup (Figure 9). This type of drawing usually elicits a question such as “what hit a red cup”. A typical third static drawing would be the same as the second drawing except the arrow depicting the hit action would indicate the red teapot as the recipient of the action (Figure 10). Another animation with a green motorcycle pushing a black cup, having only one action would have only one corresponding static drawing. In this drawing either the green motorcycle or the black cup would be replaced by a question mark and the push action would be depicted by arrows. The version of the static drawing with the motorcycle missing is shown in Figure 11.

**FIGURE 8. Static drawing to elicit “What color teapot passes through the wall?”**
FIGURE 9. Static drawing to elicit “What hit the red cup?”
FIGURE 10. Static drawing to elict “What hit the red teapot?”

FIGURE 11. Static drawing: motorcycle pushes cup, motorcycle missing
The QUESTIONER, besides asking questions from the static drawing is required to write down the “answer”, in other words the information represented by the question mark(s) on an answer sheet. The answer sheet serves the useful purpose with English speakers of providing a check on what the QUESTIONER understood, and with non-English speakers produces the additional benefit of providing at least a vocabulary list and perhaps even a rough transcription of the VIEWER’s utterances.

The Question-Answer Task is always done in the first session. In early versions of the animation elicitation technique, participants were tempted to give description at a level of detail that was unnecessary and undesirable because the requirements of the task were not as clear. Despite instructions to the contrary, participants would attempt to give exact trajectories and screen coordinates. Besides eliciting utterances that were too complicated, participants became frustrated at the lack of clear goals and the task became very long and tedious. The Question-Answer Task does not suffer from this problem because the VIEWER can have as a goal simply answering the QUESTIONER’s questions. The QUESTIONER is quite constrained in this task by the static drawings. It is usually clear to the participants by the end of the initial training session that excessive elaboration is pointless for accomplishing the task. The Question-Answer Task is also helpful in training the participants to think in terms of simple responses for the other two tasks. The major limitation of the Question-Answer Task is that sub-sentential answers to wh-questions are very natural. VIEWERS often respond to questions with noun phrases rather than entire sentences. Consequently, sentential focus phenomena are more difficult to investigate in the portion of the corpus elicited from this task.
4.4 Task 2: Directed Description

The Directed Description Task is designed to address the main limitation of the Question-Answer Task. In the Directed Description Task, the VIEWER has both the animation and the corresponding static drawings displayed. The other participant, the LISTENER, has answer sheets with a choice of abstract representations of the various actions with places to write in the characters involved. An example of the Directed Description task LISTENER answer sheet is shown in Figure 12.

FIGURE 12. Directed Description Task LISTENER answer sheet

After each action in the animation, a static drawing corresponding to that action is displayed on the VIEWER’s screen, next to the animation. The animation pauses with the display of the static drawing until the VIEWER clicks a button to continue. The VIEWER’s task is to think of the question that would be elicited by the static drawing and then describe the action by answering that question. The static drawings are the same ones used for the
Question-Answer Task. It is completely clear to the participants that they cannot give mere NP answers in this task because their partner does not know anything about the animation. The LISTENER’s task is to pick the right abstract action drawing for the action described by the VIEWER and to fill in the characters from the VIEWER’s description. Since the LISTENER’s answer sheet represents the actions in the animations quite abstractly, and since the answer sheet also limits the LISTENER to multiple choice on the actions, it helps eliminate the problem of excessive detail that was encountered in the early versions of the task. Excessive detail is also discouraged by emphasis in the instructions to the participants on the similarity of the Directed Description Task to the Question-Answer Task. The “training” provided by having done the Question-Answer task first also helps discourage excessive detail.

4.5 Task 3: Unrestricted Description

The Unrestricted Description Task is used as a sort of baseline for the Directed Description Task and the Question-Answer Task. It produces descriptions of the animations without all the constraints of the other two tasks. Discrepancies between the Unrestricted Description Task descriptions and the Directed Description Task data for the same animations can indicate that the VIEWERs in the Directed Description Task were actually attending to the drawings in formulating their descriptions. This provides some limited way to check on the efficacy of the Directed Description Task. In the Unrestricted Description Task the LISTENER has the same answer sheets as for the Directed Description Task. The VIEWER sees the animations as in the Question-Answer Task with no pauses built in. The VIEWER’s task is to describe the animations as they play so that the LISTENER can successfully fill in his or her answer sheet. Describing the action as it proceeds also helps the
VIEWER avoid memory limitations and makes the task quicker to complete. As in the Directed Description Task it is quite clear that elaborate descriptions will not be useful. Participants benefit from their experience in the previous two tasks, so that the pace of the animations keeps them attentive but is not too fast for them to handle. As in the other two tasks the LISTENER’s answer sheets provide a vocabulary list bordering on rough transcription that is particularly useful for the Hungarian data.

4.6 RELEVANT NON-UNIQUENESS

RELEVANT NON-UNIQUENESS is manipulated in the animations and drawings in two ways. It is mainipulated in the animations by having the behavior of a third character parallel the behavior of one of the characters participating in the action. For example, if a red car hits a blue rectangle the red car can be RELEVANTLY NON-UNIQUE with respect to a green car by having the green car follow a path similar to that of the red car and by having the green car veer away from hitting the blue rectangle as late as possible. The viewer is biased toward considering the cars as RELEVANTLY NON-UNIQUE with respect to each other since it plausible for either one to hit the blue rectangle during most of the animation. In this example, the viewer is also biased toward considering the red and green cars as RELEVANTLY NON-UNIQUE by virtue of them both being cars. That is, sharing a property makes objects more likely to be treated as RELEVANTLY NON-UNIQUE. In trial versions of the task RELEVANT NON-UNIQUENESS was also enhanced by using drawings in the Question-Answer Task and the Directed Description Task that elicited “which” or “either-or” questions. However this manipulation with the drawings had the major drawback of forcing INFORMATION STRUC-
TURE FOCUS and RELEVANT NON-UNIQUENESS to cooccur, and so did not provide the data needed to show the individual contributions of INFORMATION STRUCTURE and RELEVANT NON-UNIQUENESS to prosody. RELEVANT NON-UNIQUENESS is controlled in the same way in all three tasks.

4.7 INFORMATION STRUCTURE

INFORMATION STRUCTURE was manipulated through the static drawings. The “question test” is widely accepted as a test of what is the focus of an utterance, although as I discuss in Chapter 2, it has considerable limitations. It is, however, useful in question-answer pairs for diagnosing the INFORMATION STRUCTURE FOCUS in simple answers to wh-questions. By using the drawings to change which characters in an animation were questioned, the INFORMATION STRUCTURE FOCUS of the answer could be shifted in the Question-Answer Task, and the INFORMATION STRUCTURE FOCUS of the description could be shifted in the Directed Description Task. The drawings were designed so that agents and recipients of the actions were questioned in equal numbers in the course of an animation sequence.

4.8 DISCOURSE NEW/OLD

The action of having a character travel through/under a brick wall prior to the eat, hit or push actions was used to make that character DISCOURSE OLD. In the Question-Answer Task and in the Directed Description Task, a drawing is used to elicit a question about the character going through/under the wall. This virtually forces the character to be discussed
prior to the main actions, thereby making it DISCOURSE OLD. In fact, the only cases in which this strategy was occasionally unsuccessful was if there were two characters passing through the wall and the participant viewing the animation failed to attend to them both. This manipulation of DISCOURSE NEW/OLD is similar to what was done in Terken and Hirschberg (1994), except that in their task all the characters were situationally evoked for the speaker because they were all constantly present on the speaker’s screen. In the animations used to collect this corpus no character was reused in any single session, and the same shape was not used closer together than ten animations apart. This distance was selected to exceed short term memory capacity, which is only about 7 items, without neumonic assistance. Consequently, at the point a character is introduced into the animation, it is DISCOURSE NEW, hearer new (Prince 1992) and brand new (Prince 1981b). The spacing between animations was also designed to encourage participants to view each animation as separate and therefore be dissuaded from considering characters as carrying over from animation to animation.

4.9 Focus particles

Four animations designed to elicit focus particles such as just or only were included in each animation sequence. These animations had the following four “plots”:  

1. Four objects all hit a fifth object and bounce off. One of the four “hitter” objects hits the “hittee” a second time. This type of animation was an attempt to elicit something like “X,Y, Z and Q hit P, then only/just Q hit P again/a second time”
2. There are 3 or 4 identical objects of type X and 3 or 4 identical objects of type Y. There are 2 or 3 objects of type Z. Some Zs hit all the Xs and Ys, but one Z hits Xs but not Ys. This animation was designed to elicit something like “Z₁ hit all the Xs and Ys, Z₂ hit all the Xs and Ys, but Z₃ only hit the Xs”.

3. Three objects (such as scissors or axes) make contact with a fourth stationary item. Two of the three moving objects sequentially remove a piece of the fourth object. The third moving object does not remove a piece of the stationary object but instead bounces off of it. This type of animation was designed to elicit something like “Z₁ and Z₂ cut off/ate a piece of the Y, and then Z₃ just/only bounced off.

These animations were variable in their effectiveness in eliciting just or only. Speakers did use the target lexical items, but not on every presentation of the animations. Animations for eliciting even and barely were tested but found to be completely unsuccessful in inspiring the use of the desired focus particles. The “even” animations’ failure was quite intriguing. “Even” is generally thought to mark unexpectedness or improbability, but this appears to be insufficient to make “even” a natural choice for speakers. In the “even” animations appropriately sized objects went into other objects. Then unreasonably large objects went into these same objects with great apparent difficulty calculated to enhance the apparent improbability of the action. Speakers commented on how improbable the large object going into the smaller object was, but never used even.

4.10 Animation sequences

There were two animation sequences used in collecting this corpus. The same sequences were used in all three tasks with only the slight modification of additional pauses and con-
trol buttons in the animations for Task 2. The modifications for Task 2 were needed to accomplish the required difference in presentation for that task. The following table shows what actions and characters were used in each animation and where in the sequences that animation appeared. The animation scripts for animations 1 through 21 were designed to elicit utterances in which INFORMATION STRUCTURE FOCUS, RELEVANT NON-UNIQUENESS (RNU) and DISCOURSE NEW (D-NEW) do not cooccur. Since there are three factors and only a maximum of two characters per action for the “hit,” “push” and “eat” actions, two of the three factors will occur on the same character but the third factor will occur on the other character and thus be separated out. The sequences of actions for animations 1-21 in both sequences are shown in Table 1 and Table 2. These tables also show the discourse properties that each character has at each action in each animation. Table 1 gives the information for the Question-Answer task, while Table 2 provides the same information for the Directed Description Task. The discourse properties of the character for the Unrestricted Description Task will be identical to those for the Directed Description Task except that the INFORMATION STRUCTURE (IS) FOCUS will be unknown.

<table>
<thead>
<tr>
<th>ACTION</th>
<th>CHARACTER PROPERTIES</th>
<th>ORDER IN SEQUENCE</th>
</tr>
</thead>
<tbody>
<tr>
<td></td>
<td>IS</td>
<td>RNU</td>
</tr>
<tr>
<td>char1 through wall</td>
<td>c1</td>
<td></td>
</tr>
<tr>
<td>char1 ACTION char2</td>
<td>c1</td>
<td>c2</td>
</tr>
<tr>
<td>char1 ACTION char3</td>
<td>c1</td>
<td>c3</td>
</tr>
</tbody>
</table>

TABLE 1. Animation Schemata--Question-Answer Task
<table>
<thead>
<tr>
<th>ACTION</th>
<th>CHARACTER PROPERTIES</th>
<th>ORDER IN SEQUENCE</th>
</tr>
</thead>
<tbody>
<tr>
<td></td>
<td>QUESTION-ANSWER TASK</td>
<td>IS</td>
</tr>
<tr>
<td>char1, char3 through wall</td>
<td>c1, c3</td>
<td>c1, c3</td>
</tr>
<tr>
<td>char1 ACTION char2</td>
<td>c1</td>
<td>c2</td>
</tr>
<tr>
<td>char1 ACTION char3</td>
<td>c1</td>
<td>c3</td>
</tr>
<tr>
<td>char2 through wall</td>
<td>c2</td>
<td>c3</td>
</tr>
<tr>
<td>char1 ACTION char2</td>
<td>c1</td>
<td>c2</td>
</tr>
<tr>
<td>char1 ACTION char3</td>
<td>c1</td>
<td>c3</td>
</tr>
<tr>
<td>char2, char3 through wall</td>
<td>c2, c3</td>
<td>c2, c3</td>
</tr>
<tr>
<td>char1 ACTION char2</td>
<td>c1</td>
<td>c2</td>
</tr>
<tr>
<td>char1 ACTION char3</td>
<td>c1</td>
<td>c3</td>
</tr>
<tr>
<td>char1, char2 through wall</td>
<td>c1, c2</td>
<td>c1, c2</td>
</tr>
<tr>
<td>char1 ACTION char2</td>
<td>c1</td>
<td>c2</td>
</tr>
<tr>
<td>char ACTION char3</td>
<td>c1</td>
<td>c3</td>
</tr>
<tr>
<td>char1 ACTION char2</td>
<td>c1</td>
<td>c2</td>
</tr>
<tr>
<td>char1 ACTION char3</td>
<td>c1</td>
<td>c3</td>
</tr>
<tr>
<td>char3 through wall</td>
<td>c3</td>
<td>c3</td>
</tr>
<tr>
<td>char1 ACTION char2</td>
<td>c1</td>
<td>c2</td>
</tr>
<tr>
<td>char1 ACTION char3</td>
<td>c1</td>
<td>c3</td>
</tr>
<tr>
<td>char1 through wall</td>
<td>c1</td>
<td>c1</td>
</tr>
<tr>
<td>char1 ACTION char2</td>
<td>c2</td>
<td>c1</td>
</tr>
<tr>
<td>char3 ACTION char2</td>
<td>c2</td>
<td>c3</td>
</tr>
<tr>
<td>ACTION</td>
<td>CHARACTER PROPERTIES</td>
<td>ORDER IN SEQUENCE</td>
</tr>
<tr>
<td>-------------------</td>
<td>----------------------</td>
<td>-------------------</td>
</tr>
<tr>
<td></td>
<td>IS</td>
<td>RNU</td>
</tr>
<tr>
<td>char1, char3 through wall</td>
<td>c1,c3</td>
<td></td>
</tr>
<tr>
<td>char1 ACTION char2</td>
<td>c2</td>
<td>c1</td>
</tr>
<tr>
<td>char3 ACTION char2</td>
<td>c2</td>
<td>c3</td>
</tr>
<tr>
<td>char2 through wall</td>
<td>c2</td>
<td></td>
</tr>
<tr>
<td>char1 ACTION char2</td>
<td>c2</td>
<td>c1</td>
</tr>
<tr>
<td>char3 ACTION char2</td>
<td>c2</td>
<td>c3</td>
</tr>
<tr>
<td>char2, char3 through wall</td>
<td>c2,c3</td>
<td></td>
</tr>
<tr>
<td>char1 ACTION char2</td>
<td>c2</td>
<td>c1</td>
</tr>
<tr>
<td>char3 ACTION char2</td>
<td>c2</td>
<td>c3</td>
</tr>
<tr>
<td>char1, char2 through wall</td>
<td>c1,c2</td>
<td></td>
</tr>
<tr>
<td>char1 ACTION char2</td>
<td>c2</td>
<td>c1</td>
</tr>
<tr>
<td>char3 ACTION char2</td>
<td>c2</td>
<td>c3</td>
</tr>
<tr>
<td>char1 ACTION char2</td>
<td>c2</td>
<td>c1</td>
</tr>
<tr>
<td>char3 ACTION char2</td>
<td>c2</td>
<td>c3</td>
</tr>
<tr>
<td>char3 through wall</td>
<td>c3</td>
<td></td>
</tr>
<tr>
<td>char1 ACTION char2</td>
<td>c2</td>
<td>c1</td>
</tr>
<tr>
<td>char3 ACTION char2</td>
<td>c2</td>
<td>c3</td>
</tr>
<tr>
<td>char1 ACTION char2</td>
<td>c1</td>
<td></td>
</tr>
<tr>
<td>char1 ACTION char2</td>
<td>c2</td>
<td>c1</td>
</tr>
</tbody>
</table>
### TABLE 1. Animation Schemata--Question-Answer Task (Continued)

<table>
<thead>
<tr>
<th>ACTION</th>
<th>CHARACTER PROPERTIES</th>
<th>ORDER IN SEQUENCE</th>
</tr>
</thead>
<tbody>
<tr>
<td></td>
<td>IS</td>
<td>RNU</td>
</tr>
<tr>
<td>char1 through wall</td>
<td>c1</td>
<td></td>
</tr>
<tr>
<td>char1 ACTION char2</td>
<td>c1</td>
<td></td>
</tr>
<tr>
<td>char2 through wall</td>
<td>c2</td>
<td></td>
</tr>
<tr>
<td>char1 ACTION char2</td>
<td>c2</td>
<td></td>
</tr>
<tr>
<td>char1 ACTION char2</td>
<td>c2</td>
<td></td>
</tr>
</tbody>
</table>

### TABLE 2. Animation Schemata--Directed Description Task

<table>
<thead>
<tr>
<th>ACTION</th>
<th>CHARACTER PROPERTIES</th>
<th>ORDER IN SEQUENCE</th>
</tr>
</thead>
<tbody>
<tr>
<td></td>
<td>INFO</td>
<td>CONT</td>
</tr>
<tr>
<td>char1 through wall</td>
<td>c1</td>
<td></td>
</tr>
<tr>
<td>char1 ACTION char2</td>
<td>c1</td>
<td>c2</td>
</tr>
<tr>
<td>char1 ACTION char3</td>
<td>c1</td>
<td>c3</td>
</tr>
<tr>
<td>char1, char3 through wall</td>
<td>c1,c3</td>
<td></td>
</tr>
<tr>
<td>char1 ACTION char3</td>
<td>c1</td>
<td>c3</td>
</tr>
<tr>
<td>char2 through wall</td>
<td>c2</td>
<td></td>
</tr>
<tr>
<td>char1 ACTION char2</td>
<td>c1</td>
<td>c2</td>
</tr>
<tr>
<td>char1 ACTION char3</td>
<td>c1</td>
<td>c3</td>
</tr>
<tr>
<td>char2, char3 through wall</td>
<td>c2,c3</td>
<td></td>
</tr>
</tbody>
</table>

70
TABLE 2. Animation Schemata--Directed Description Task (Continued)

<table>
<thead>
<tr>
<th>ACTION</th>
<th>CHARACTER PROPERTIES</th>
<th>ORDER IN SEQUENCE</th>
</tr>
</thead>
<tbody>
<tr>
<td></td>
<td>DIRECTED DESCRIPTION TASK</td>
<td>INFO</td>
</tr>
<tr>
<td>char1 ACTION char2</td>
<td>c1</td>
<td>c2</td>
</tr>
<tr>
<td>char1 ACTION char3</td>
<td>c1</td>
<td>c3</td>
</tr>
<tr>
<td>char1, char2 through wall</td>
<td>c1,c2</td>
<td></td>
</tr>
<tr>
<td>char1 ACTION char2</td>
<td>c1</td>
<td>c2</td>
</tr>
<tr>
<td>char1 ACTION char3</td>
<td>c1</td>
<td>c3</td>
</tr>
<tr>
<td>char1 ACTION char2</td>
<td>c1</td>
<td>c2</td>
</tr>
<tr>
<td>char1 ACTION char3</td>
<td>c1</td>
<td>c3</td>
</tr>
<tr>
<td>char3 through wall</td>
<td></td>
<td>c3</td>
</tr>
<tr>
<td>char1 ACTION char2</td>
<td>c1</td>
<td>c2</td>
</tr>
<tr>
<td>char1 ACTION char3</td>
<td>c1</td>
<td>c3</td>
</tr>
<tr>
<td>char1 through wall</td>
<td>c1</td>
<td></td>
</tr>
<tr>
<td>char1 ACTION char2</td>
<td>c2</td>
<td>c1</td>
</tr>
<tr>
<td>char3 ACTION char2</td>
<td>c2</td>
<td>c3</td>
</tr>
<tr>
<td>char1, char3 through wall</td>
<td>c1,c3</td>
<td></td>
</tr>
<tr>
<td>char1 ACTION char2</td>
<td>c2</td>
<td>c1</td>
</tr>
<tr>
<td>char3 ACTION char2</td>
<td>c2</td>
<td>c3</td>
</tr>
<tr>
<td>char2 through wall</td>
<td>c2</td>
<td></td>
</tr>
<tr>
<td>char1 ACTION char2</td>
<td>c2</td>
<td>c1</td>
</tr>
<tr>
<td>char3 ACTION char2</td>
<td>c2</td>
<td>c3</td>
</tr>
</tbody>
</table>
Animations 22-25 in each sequence were of the type discussed in section 2.10 which were
designed to elicit focus particles.

4.11 Training

Training is very useful in this task because it gives the participants an opportunity to develop simple ways of talking about the actions and makes them comfortable with the task. Untrained participants can find the task quite difficult and tend to provide an excruciating level of detail which makes examining the effect of the discourse/semantic factors of interest more difficult. Also the level of detail and task difficulty problems observed with untrained participants persists even with small amounts of training for the Directed Description and Unrestricted Description Tasks. The problem is elevated by having participants do the tasks in the following order: Question-Answer, Directed Description, Unrestricted Description. The more constrained tasks serve as training for the less constrained ones. Training prior to the Question-Answer task consists of three parts:

1. Naming actions. Participants are shown animations and drawings and asked to supply verbs or verb phrase “names” for the actions.

2. Cooperative Previews. Pairs of task partners watch displays that pair animations with corresponding diagrammatic drawings. They are encouraged to discuss how the drawings correspond to the animations, how to describe the actions, and to practice their task.

3. Practice Runs.
Prior to performing the Question-Answer Task the participants watched eight animation-drawing pairs together. After viewing the previews, participants were given four animation-drawing sets to do for practice. Written instructions were given to participants as part of the preview and practice sequences and the investigator was on hand to answer questions about the task during the training period. The same training procedure was also used for the Directed Description Task. The Unrestricted Description task generally required no training beyond what has already been acquired by participants in the course of doing the Question-Answer and Directed Description Tasks.

4.12 The "Jeopardy Task"

As previously discussed, the Unrestricted Description Task has no manipulation of INFORMATION STRUCTURE status built into the task. Consequently, there is no reliable evidence from the task itself about which parts of an utterance are INFORMATION STRUCTURE FOCUS or INFORMATION STRUCTURE GROUND. To address this problem, the Jeopardy Task was designed to elicit judgements from subjects about the INFORMATION STRUCTURE status of an utterance. The Jeopardy Task is named after a popular television game show in which contestants are given an answer by the host and must respond with the matching question in order to win prizes. The Jeopardy Task has a similar format, but differs from the Jeopardy game show in giving its contestants/subjects the advantage of multiple choice and the disadvantage of much less impressive prizes. In the Jeopardy Task, the subjects are the contestants and the host, either a computer interface or an audio tape, supplies answers from the Unrestricted Description Task data. As an example, if an utterance/answer “The blue bicycle hit the yellow tree” was played to subjects, they would be given a choice among
“What hit the yellow tree?” which would indicate that “the blue bicycle” is the INFORMATION STRUCTURE FOCUS, “What did the blue bicycle hit?” which corresponds to having “the yellow tree” as the INFORMATION STRUCTURE FOCUS and “What happened?” which would indicate that the whole utterance is INFORMATION STRUCTURE FOCUS. Coding for INFORMATION STRUCTURE was based on a match between two of three jeopardy participants. Cases for which there was no agreement among the jeopardy participants were not used in the analysis.

The Jeopardy Task was also useful as a check on the Directed Description Task data. It was clear that some VIEWERS in the Directed Description Task were more diligent than others in attempting to answer the question prompted by the static drawing.
Chapter 5

English

This chapter discusses the analysis of the English Animation Corpus and the results of that analysis. Since prosody is particularly important to focus realization in English this analysis concentrates on investigating various aspects of prosody associated with the focus interpretations identified in Chapters 2 and 3. There are three major parts to the analysis. Section 5.1 is concerned with duration and amplitude effects for INFORMATION STRUCTURE FOCUS, DISCOURSE NEW/OLD and RELEVANT NON-UNIQUENESS across all the usable data. Only INFORMATION STRUCTURE is found to have a strong correlation with duration and amplitude.

Section 5.2 investigates how the greater duration and amplitude shown in Section 5.1 are distributed within INFORMATION STRUCTURE foci. Whether duration and amplitude effects in a multi-word focus constituent reside on a single word or appear on all the components of a focus constituent is a direct test of the prevalent focus projection approach. This type of account takes pitch accent as the sole means of expressing focus and maintains that a pitch accent on a word ambiguously marks that word or larger constituents containing that word as focus. Duration and amplitude effects are found on all components of INFORMATION STRUCTURE FOCUS NPs in the English Animation Corpus. This suggests that duration and amplitude (and whatever they represent phonologically) also mark INFORMATION
STRUCTURE FOCUS. Unlike pitch accent it appears that duration and amplitude mark the extent of the INFORMATION STRUCTURE FOCUS constituent directly and unambiguously.

Finally, Section 5.3 presents an analysis of F0 effects for INFORMATION STRUCTURE FOCUS, DISCOURSE NEW and RELEVANT NON-UNIQUENESS. As with the result of the duration and amplitude analysis, F0 effects are consistently associated only with INFORMATION STRUCTURE FOCUS.

5.1 Duration and Amplitude

This Section examines duration and amplitude in relation to the three notions of “focus” discussed in the previous chapters: INFORMATION STRUCTURE, DISCOURSE NEW and RELEVANT NON-UNIQUENESS. The data used is from the Animation Corpus discussed in Chapter 4. Matched pairs of words are selected from the corpus to test whether INFORMATION STRUCTURE FOCUS (IS-focus) has an effect on duration or amplitude compared to INFORMATION STRUCTURE GROUND (IS-ground), whether being DISCOURSE NEW has an effect on duration and amplitude compared to being DISCOURSE OLD, and whether RELEVANTLY NON-UNIQUE items differ in duration and amplitude from those that are not. The same pairs are used for investigating both duration and amplitude. In selecting matched pairs two tokens of the same lexical item are chosen that share all relevant properties except the one being considered. For example, triangle in (56) is paired with triangle in (57).

(56)

136:B: okay

137:A: what goes through the divider?
138:B: a black helicopter and a red triangle

139:A: and a red triangle

140:B: uh-huh

141:A: okay

142:A: uh what does the black helicopter push?

143:B: it pushes the red triangle

(57)

1:A: okay

2:B: okay

3:A: alright what color is the apple?

4:B: the apple is red

5:A: okay

6:B: standard color for an apple

7:A: what hits the triangle the yellow triangle

8:B: a red apple hits the yellow triangle

Both are instances of the lexical item triangle, both are DISCOURSE OLD, and not RELEVANTLY NON-UNIQUE and thus matched for those properties as well as for speaker and task. The token of triangle in (56) is INFORMATION STRUCTURE FOCUS and the one in (57) is INFORMATION STRUCTURE GROUND.

The relevant measurement or calculation (e.g. duration, amplitude or amplitude integral) is taken on both items of each pair and their values are plotted as the X and Y coordinates of
points in the scatter plots shown later in this chapter. The difference between the members of a matched pair in the value of the relevant measurement or calculation is used to test statistically whether the two types of items represented by the matched pairs significantly differ from each other. This approach using matched pairs allows control of segmental and other known or likely effects on prosody. Since different segments have different inherent lengths (Lehiste 1970) it is important to match the segmental material or the segmental effects could overwhelm an effect due to the factor of interest. The position of the token with respect to the intonational and intermediate phrase boundaries must also be controlled in the matched pairings, since it is well known that syllables immediately preceding phrase boundaries show phrase final lengthening in English. Pairs are also matched for speaker, task and as closely as possible for linear order (i.e. match word #4 of one utterance to word #5 rather than word #1 of another utterance). The integral of amplitude across duration is also calculated for the same matched pairs as are used for amplitude and duration. This amplitude integral can be seen as representing overall energy. It would be expected to show effects that might not appear in either the duration or amplitude alone if speakers have a choice of increasing either duration or amplitude to mark prominence or if the combination of the two reaching some threshold is a way of achieving prominence. The statistic used in testing the significance of differences between INFORMATION STRUCTURE FOCUS/INFORMATION STRUCTURE GROUND, DISCOURSE NEW/ DISCOURSE OLD and RELEVANT NON-UNIQUENESS/not RELEVANT NON-UNIQUENESS is the Wilcoxon Signed-Rank Test. This test is non-parametric. Since the distributions are rarely normal, the more obvious paired t-test, which requires normal distributions, is unreliable for this data. The Wilcoxon Signed-Rank Test also has the useful property of being a
valid test when the number of items is as small as 6. Table 3 shows the null and alternative hypotheses addressed by the test. The null hypothesis is rejected if the appropriate statistic is less than or equal to a critical value. Examples of some critical values are given in Table 4.

### TABLE 3. Wilcoxon Signed-Rank Test - hypotheses and statistics

<table>
<thead>
<tr>
<th>To Test</th>
<th>Versus</th>
<th>Compute</th>
</tr>
</thead>
<tbody>
<tr>
<td>Null Hypothesis $(H_0)$</td>
<td>Alternative Hypothesis $(H_1)$</td>
<td></td>
</tr>
<tr>
<td>$\mu_1 = \mu_2$</td>
<td>$\mu_1 &lt; \mu_2$</td>
<td>$w_+$</td>
</tr>
<tr>
<td></td>
<td>$\mu_1 &gt; \mu_2$</td>
<td>$w_-$</td>
</tr>
<tr>
<td></td>
<td>$\mu_1 \neq \mu_2$</td>
<td>$w$</td>
</tr>
</tbody>
</table>

### TABLE 4. Critical Values for the Wilcoxon Signed-Rank Test

<table>
<thead>
<tr>
<th>$n$</th>
<th>One-sided $\alpha = 0.01$</th>
<th>One-sided $\alpha = 0.025$</th>
<th>One-sided $\alpha = 0.05$</th>
<th>Two-sided $\alpha = 0.02$</th>
<th>Two-sided $\alpha = 0.05$</th>
<th>Two-sided $\alpha = 0.10$</th>
</tr>
</thead>
<tbody>
<tr>
<td>5</td>
<td>1</td>
<td>1</td>
<td>1</td>
<td>1</td>
<td>1</td>
<td>1</td>
</tr>
<tr>
<td>6</td>
<td>1</td>
<td>2</td>
<td>2</td>
<td>2</td>
<td>2</td>
<td>2</td>
</tr>
<tr>
<td>7</td>
<td>2</td>
<td>2</td>
<td>4</td>
<td>4</td>
<td>4</td>
<td>4</td>
</tr>
<tr>
<td>8</td>
<td>4</td>
<td>4</td>
<td>6</td>
<td>6</td>
<td>6</td>
<td>6</td>
</tr>
<tr>
<td>9</td>
<td>6</td>
<td>6</td>
<td>8</td>
<td>8</td>
<td>8</td>
<td>8</td>
</tr>
<tr>
<td>10</td>
<td>8</td>
<td>8</td>
<td>11</td>
<td>11</td>
<td>11</td>
<td>11</td>
</tr>
</tbody>
</table>

#### 5.1.1 Duration

Examples (58) and (59) show a matched pair for the INFORMATION STRUCTURE FOCUS/INFORMATION STRUCTURE GROUND distinction. The context shows that the underlined tokens of the word *red* in the two examples are both DISCOURSE OLD and RELEVANTLY NON-UNIQUE. Recall that the diagnostic for INFORMATION STRUCTURE FOCUS is being the
minimal answer to a wh-question. The token of red in (58) is coded as INFORMATION STRUCTURE FOCUS since it is part of the constituent the red mop, which is the minimal answer to the immediately preceding wh-question: what pushes the white box? The token of red in (59) on the other hand is coded as INFORMATION STRUCTURE GROUND since it is not part of the minimal answer to the immediately preceding wh-question: okay what does the red shoe hit?.

(58)
110:B: okay
111:A: uh what goes past the divider
112:B: a white box and a red mop and
113:A: too
114:B: a blue envelope
115:A: a blue what?
116:B: envelope
117:A: envelope
118:A: okay and all of them are on the first side of the divider?
119:B: yes and they move to the
120:A: and they all move to the left
121:B: other side
122:A: or to the right I'm sorry
123:A: okay
124:A: what pushes the white box
125:B: the red mop pushes the white box
31: B: okay

32: A: what color is the shoe?

33: B: there are two shoes

34: A: okay before it hits the divider or after it hits the divider?

35: B: um one shoe comes down behind the divider and as it emerges from the other side another shoe comes down at the same time

36: A: okay what color is the first shoe?

37: B: the first shoe is red

38: A: okay

39: A: okay what does the red shoe hit?

40: B: the red shoe hits a blue teapot

The waveform displays of the red mop pushes the white box in Figure 13 and of the red shoe hits a blue teapot in Figure 14 show the durations of the two tokens of red as “D:<number>” in the shaded part of the top bar of each display. The INFORMATION STRUCTURE FOCUS red from (58) has duration of 0.24875 seconds while the INFORMATION STRUCTURE GROUND red from (59) has duration of 0.22612 seconds. The shaded portion of the waveform in each display indicates the portion of waveform associated with the item red.
FIGURE 13. Waveform for *the red mop pushes the white box showing duration of red*

FIGURE 14. Waveform for *the red shoe hits a blue teapot showing duration of red*

The scatterplot in Figure 15 shows the effect of INFORMATION STRUCTURE FOCUS versus INFORMATION STRUCTURE GROUND on duration. The duration of the INFORMATION STRUCTURE FOCUS (IS-focus) member of each matched pair is plotted on the Y-axis and the duration of the INFORMATION STRUCTURE GROUND (IS-ground) member of the matched pair is plotted on the X-axis. For example, the duration of the token of *red* from (58) is plotted on the Y-axis against the duration of the token of *red* from (59) on the X-axis to produce the point labeled with the “x” in Figure 15.
It is clear from visual inspection that INFORMATION STRUCTURE FOCUS (IS-focus) items tend to be longer than their INFORMATION STRUCTURE GROUND (IS-ground) counterparts.

The $x=y$ line plots where the durations of INFORMATION STRUCTURE FOCUS and INFORMATION STRUCTURE GROUND are equal.
TION STRUCTURE GROUND members of a matched pair would be equal. In Figure 15 most of the points fall above the x=y line in the region where INFORMATION STRUCTURE FOCUS duration is longer than INFORMATION STRUCTURE GROUND duration. In addition, the points that do fall below x=y line are relatively close to the x=y line, particularly in comparison to the points that fall above it. Since the Wilcoxon Signed-Rank test takes into account both direction and magnitude, it is not surprising that the results show that the duration of INFORMATION STRUCTURE FOCUS items is greater than the duration of INFORMATION STRUCTURE GROUND items at a significance level of 0.003.

An example of a matched pair for DISCOURSE NEW versus DISCOURSE OLD is given in (60) and (61). The tokens of green in (60) and (61) are both INFORMATION STRUCTURE FOCUS and not RELEVANT NON-UNIQUENESS. The token of green in (60) is DISCOURSE NEW, while the token of green in (61) is DISCOURSE OLD.

(60)

207:B:okay
208:A:what crosses the divider?
209:B:a green lock crosses the divider

(61)

68:B:okay
69:A:uh what color is the anchor
70:B:the anchor is green
71:A:and what does the paperclip hit?
72:B:the paperclip hits the green anchor
The waveform display for *a green lock crosses the divider* is shown in Figure 16 and *the paperclip hits the green anchor* in Figure 17. The duration of the DISCOURSE NEW token of *green* from (60) can be seen from the shaded duration indicator in Figure 16 to be 0.42538 seconds. The duration of the DISCOURSE OLD member of this matched pair can be seen from the shaded duration indicator in Figure 17 to be 0.35800 seconds.

**FIGURE 16. waveform for *a green lock crosses the divider* showing duration of *green***

**FIGURE 17. waveform for *the paperclip hits the green anchor* showing duration of *green***

The scatterplot in Figure 18 shows that there is no effect for being DISCOURSE NEW versus being DISCOURSE OLD. The durations of DISCOURSE NEW items are plotted on the Y-axis and the durations of their DISCOURSE OLD partners are plotted on the X-axis. For example, the duration of 0.40538 seconds for the DISCOURSE NEW token of *green* in (60) is plotted
on the Y-axis against the 0.35800 second duration of the DISCOURSE OLD token of green in (61) on the X-axis, to produce the point marked by the “x” in Figure 18.

**FIGURE 18. Matched Pairs: DISCOURSE NEW vs OLD durations**

**Test Mean (Dnew - Dold) =value**

Hypothesized Value 0
Actual Estimate 0.03433

**Wilcoxon Signed-Rank Test**

Test Statistic 95.500
Prob > |t| 0.302
The points in Figure 18 appear to be clustered fairly evenly on both sides of the x=y line. This distribution indicates that there is no consistent tendency for either DISCOURSE NEW or DISCOURSE OLD items to have greater duration. This appraisal of the scatter plot in Figure 18 is confirmed by the results of the Wilcoxon Signed-Rank Test which show no significant difference in either direction. The conclusion is that there is no effect on duration associated with being DISCOURSE NEW or DISCOURSE OLD.

Examples (62) and (63) show a matched pair for RELEVANT NON-UNIQUENESS versus not RELEVANT NON-UNIQUENESS. Both examples have tokens of the item white that are both INFORMATION STRUCTURE GROUND and DISCOURSE OLD. The instance of white in (62) is RELEVANTLY NON-UNIQUE; the white box forms a contrast set with the blue envelope since they both underwent the action of going past the divider. At the point in (62) where white is uttered, it alone allows reduction of the possible referents, white box and blue envelope, to just the white box. The instance of white in (63) is not RELEVANTLY NON-UNIQUE.

(62)

110:B: okay
111:A: uh what goes past the divider
112:B: a white box and a red mop and
113:A: too
114:B: a blue envelope
115:A: a blue what?
116:B: envelope
117:A: envelope
118:A: okay and all of them are on the first side of the divider?
119:B: yes and they move to the
120:A: and they all move to the left
121:B: other side
122:A: or to the right I'm sorry
123:A: okay
124:A: what pushes the white box
125:B: **the red mop pushes the white box**

22:B: okay
23:A: what color's the oval?
24:B: it's blue
25:A: okay
26:A: what hits the white bed?
27:B: **the blue oval hits the white bed**
The duration of the RELEVANTLY NON-UNIQUE instance of white in (62) can be seen in the shaded duration indicator of Figure 19 to be 0.26875 seconds, while the duration of the not RELEVANTLY NON-UNIQUE instance of white in this matched pair can be seen from Figure 20 to have a duration of 0.24950 seconds.

In the scatter plot in Figure 21 the duration of the RELEVANTLY NON-UNIQUE (RNU) member of each matched pair is plotted on the Y-axis against the duration of the not RELEVANTLY NON-UNIQUE (not RNU) member of the matched pair on the X-axis. The point which represents the durations of the tokens of white in (62) and (63) is represented in Figure 21 by an “x”.

FIGURE 19. waveform for *the red mop pushes the white box showing duration of white*

FIGURE 20. waveform for *the blue oval hits the white bed showing duration of white*
Like the results for DISCOURSE OLD versus DISCOURSE NEW, there is no effect on duration from being RELEVANTLY NON-UNIQUE or not. Inspection of the scatter plot in Figure 21
shows a fairly even distribution of points around the x=y line, indicating no clear tendency in either direction. This is supported by the Wilcoxon Signed-Rank Test which shows no significant difference in any direction. The conclusion from this evidence is that RELEVANT NON-UNIQUENESS is not responsible for any effect on duration.

Table 5 gives a summary of the results for duration. For duration, at least, only INFORMATION STRUCTURE has an effect when the data is controlled so that the discourse factors can be investigated individually. There is no evidence of a consistent effect of either RELEVANT NON-UNIQUENESS or DISCOURSE NEW/OLD. The conclusion for duration is that INFORMATION STRUCTURE FOCUS items are consistently longer than INFORMATION STRUCTURE GROUND items, and that RELEVANTLY NON-UNIQUE and DISCOURSE NEW items that might have been expected to have similar effects based on previous literature, do not.

TABLE 5. Summary of Effect of Discourse Factors on Duration

<table>
<thead>
<tr>
<th>Discourse Factor</th>
<th>Duration (significance level)</th>
</tr>
</thead>
<tbody>
<tr>
<td>INFORMATION STRUCTURE FOCUS</td>
<td>&gt;IS-ground (0.003)</td>
</tr>
<tr>
<td>INFORMATION STRUCTURE GROUND</td>
<td>&lt;IS-focus (0.003)</td>
</tr>
<tr>
<td>DISCOURSE NEW</td>
<td>no effect (0.151)</td>
</tr>
<tr>
<td>DISCOURSE OLD</td>
<td>no effect (0.151)</td>
</tr>
<tr>
<td>RELEVANT NON-UNIQUENESS</td>
<td>no effect (0.350)</td>
</tr>
<tr>
<td>not RELEVANT NON-UNIQUENESS</td>
<td>no effect (0.350)</td>
</tr>
</tbody>
</table>
5.1.2 Amplitude

One issue in measuring amplitude is whether to use mean amplitude across the word or whether to look at maximum amplitude. Maximum amplitude is a good measure if one believes that peaks in amplitude, that is reaching some level of loudness sometime during a word, is most perceptually relevant. On the other hand, mean amplitude would be good if one believes that the overall loudness of a word is the thing which is perceptually salient. The mean could obliterate the difference between a word that peaks at quite a high amplitude and another word which maintains a steady but much lower amplitude throughout, giving somewhat counterintuitive results in some cases. However, measuring maximum amplitude could also be counterintuitive in some cases since it only considers one point in the word. Fortunately, it is not really necessary to choose between these measures since for the data examined in this study their results do not conflict. The compatibility of results for maximum and mean amplitudes in this study will be clear from the discussion in this section.

As was mentioned earlier, the same matched pairs are used for amplitude as were used for duration. For continuity, the same example matched pairs will be used for the discussion of amplitude as were used in the previous section on duration. In this section plots of amplitude for the matched pair lexical items accompany the waveform display of the sentence in which each lexical item occurred.

The waveform displays in Figure 22 and Figure 23 show the waveforms and associated amplitude plots for INFORMATION STRUCTURE FOCUS and INFORMATION STRUCTURE GROUND tokens of red (for transcripts of the preceding context see (55) and (56)). The
mean amplitudes are indicated on the amplitude plots by grey lines. As can be seen from the amplitude plot in Figure 22, the mean amplitude value for the INFORMATION STRUCTURE FOCUS token of red is 3218.647. The mean amplitude for its INFORMATION STRUCTURE GROUND partner can be seen from Figure 23 to be 3022.471.

**FIGURE 22.** Waveform for the red mop pushes the white box with amplitude plot for red showing mean amplitude
FIGURE 23. Waveform for the red shoe hits a blue teapot with amplitude plot for red showing mean amplitude.

Figure 24 is the scatterplot of the mean amplitude values for INFORMATION STRUCTURE FOCUS versus INFORMATION STRUCTURE GROUND matched pairs. The mean amplitude of the INFORMATION STRUCTURE FOCUS member of each matched pair is plotted as the Y-coordinate and the mean amplitude of the INFORMATION STRUCTURE GROUND member of the matched pair is plotted as the X-coordinate. For example, the mean amplitude value of 3218.647 for the INFORMATION STRUCTURE FOCUS (IS-focus) token of red in Figure 22 is the Y-coordinate and the mean amplitude value of 3022.647 for INFORMATION STRUCTURE
GROUND (IS-ground) token of red in Figure 23 is the X-coordinate of the point indicated on the Figure 24 scatterplot by an “x”.

**FIGURE 24. Matched Pairs: INFORMATION STRUCTURE FOCUS vs GROUND Mean Amplitude**

![Graph showing matched pairs of IS-focus(mean) vs IS-ground(mean). The x = y line is plotted.]  

**Test Mean (ISfocus - ISground) =**value

Hypothesized Value 0  
Actual Estimate 756.312

**Wilcoxon Signed-Rank Test**

<table>
<thead>
<tr>
<th>Statistic</th>
<th>Value</th>
</tr>
</thead>
<tbody>
<tr>
<td>Test Statistic</td>
<td>131.000</td>
</tr>
<tr>
<td>Prob &gt;</td>
<td>t</td>
</tr>
<tr>
<td>Prob &gt; t</td>
<td>0.006</td>
</tr>
<tr>
<td>Prob &lt; t</td>
<td>0.994</td>
</tr>
</tbody>
</table>
Visual inspection of the Figure 24 scatterplot shows that more of the points fall in the area above the x=y line, where the INFORMATION STRUCTURE FOCUS mean amplitude exceeds the INFORMATION STRUCTURE GROUND mean amplitude for a matched pair. In addition, the points that fall above the x=y line are generally further away from the line than the points falling below. This indicates that even for the few points where the mean amplitude of the INFORMATION STRUCTURE GROUND member of a pair is bigger, it is generally not very much bigger than the mean amplitude of its INFORMATION STRUCTURE FOCUS counterpart. The results of the Wilcoxon Signed-Rank Test confirm this assessment of the scatterplot. The test shows that mean amplitude for INFORMATION STRUCTURE FOCUS differs from the mean amplitude of INFORMATION STRUCTURE GROUND at a p<0.012 level of significance using the two tailed version of the test. The test results for the one tailed test lead to the conclusion that at a p<0.006 level of significance, mean amplitude of INFORMATION STRUCTURE FOCUS items is bigger than mean amplitude of INFORMATION STRUCTURE GROUND items.

Moving on to mean amplitude for DISCOURSE NEW versus DISCOURSE OLD items, waveform displays and amplitude plots are given in Figure 25 and Figure 26 for the example matched pair of the lexical item *green*. Recalling the discussion of this matched pair from the previous section, both tokens are INFORMATION STRUCTURE FOCUS and not RELEVANTLY NON-UNIQUE (the context for these examples is available in 57 and 58 of the preceding section). Figure 25 shows that the mean amplitude for the DISCOURSE NEW token of *green*, as indicated by the gray line on the amplitude plot, is 2251.063. The mean ampli-
tude of the DISCOURSE OLD token of *green*, from the amplitude plot in Figure 26, is 1293.389.

**FIGURE 25.** Waveform for a green lock crosses the divider with amplitude plot for green showing mean amplitude
FIGURE 26. Waveform for the paperclip hits the green anchor with amplitude plot for green showing mean amplitude

In the scatterplot in Figure 27, points are plotted from a set of matched pairs similar to the example matched pair of tokens of green examined in some detail in Figure 25 and Figure 26. For each matched pair the mean amplitude of the DISCOURSE NEW member is plotted as the Y-coordinate of a point while the mean amplitude of the DISCOURSE OLD member of the pair is plotted as the point’s X-coordinate. For example, the point indicated
on the Figure 27 scatterplot by an “x” has as its Y-coordinate the mean amplitude value 2251.063 of the DISCOURSE NEW token of green in Figure 25. The X-coordinate of this point, 1293.389, is the mean amplitude value of the DISCOURSE OLD token of green in Figure 26.

FIGURE 27. Matched Pairs: DISCOURSE NEW vs OLD Mean Amplitude

Test Mean (Dnew - Dold) = value

Hypothesized Value 0
Actual Estimate 100.259
The points in the Figure 27 scatterplot are fairly evenly distributed around the x=y line with no obvious trend toward either DISCOURSE NEW or DISCOURSE OLD items being consistently of greater amplitude than their counterparts of the other type. This interpretation of the scatterplot is supported by the results of the Wilcoxon Signed-Rank Test, which shows no significance for either one or two tailed versions of the test. From this data it would be reasonable to conclude that the DISCOURSE NEW/DISCOURSE OLD distinction has no effect on mean amplitude.

Figure 28 and Figure 29 show the waveform displays and amplitude plots for the example RELEVANTLY NON-UNIQUE/not RELEVANTLY NON-UNIQUE matched pair of the lexical item *white*. As with the examples for INFORMATION STRUCTURE FOCUS versus INFORMATION STRUCTURE GROUND and DISCOURSE NEW versus DISCOURSE OLD, the context for this matched pair is given in the previous section (examples 59 and 60). As previously discussed, both instances of *white* are INFORMATION STRUCTURE GROUND and DISCOURSE OLD. The mean amplitude of the RELEVANTLY NON-UNIQUE member of the pair shown in Figure 28 is 1209.999. The mean amplitude of the not RELEVANTLY NON-UNIQUE member of the pair as shown in Figure 29 is 1176.0907.
FIGURE 28. Waveform for the red mop pushes the white box with amplitude plot for white showing mean amplitude.
The scatterplot in Figure 30 displays points which have as their Y-coordinate the mean amplitude of the RELEVANTLY NON-UNIQUE member of a matched pair and have the mean amplitude of the not RELEVANTLY NON-UNIQUE member as their X-coordinate. The mean amplitude values of 1209.999 for the RELEVANTLY NON-UNIQUE (RNU) token of *white*
and $1176.0907$ for the not RELEVANTLY NON-UNIQUE (not RNU) token of white from the example matched pair are plotted in Figure 30 as the “x”.

**FIGURE 30. Matched Pairs: RELEVANTLY NON-UNIQUE vs NOT Mean Amplitude**

**Test Mean (RNU - not RNU) =value**

<table>
<thead>
<tr>
<th>Hypothesized Value</th>
<th>Actual Estimate</th>
</tr>
</thead>
<tbody>
<tr>
<td>0</td>
<td>-488.07</td>
</tr>
</tbody>
</table>

**Wilcoxon Signed-Rank Test**

| Test Statistic   | Prob > |t|   | Prob > t | Prob < t |
|------------------|--------|-----|----------|----------|
| -10.000          | 0.413  | 0.793 | 0.207    |
In Figure 30, the points are fairly evenly clustered around the x=y line, and most of the points are fairly close to the line. This would not seem to indicate any consistent difference between RELEVANTLY NON-UNIQUE and not RELEVANTLY NON-UNIQUE items in relation to mean amplitude. In fact, the Wilcoxon Signed-Rank test is not significant for either one or two tailed versions, also supporting this conclusion. As was the case for duration, the RELEVANT NON-UNIQUENESS versus not RELEVANT NON-UNIQUENESS distinction shows no effect on mean amplitude.

Next, consider maximum amplitude. Figure 31 and Figure 32 show waveform displays accompanied by amplitude plots for the INFORMATION STRUCTURE FOCUS versus INFORMATION STRUCTURE GROUND matched pair of tokens of red. These figures are similar to those shown previously for mean amplitude, except that the maximum amplitude is indicated on the amplitude plot instead of the mean amplitude. From Figure 31, the maximum amplitude of the INFORMATION STRUCTURE FOCUS token of red is 6464.66, while the maximum amplitude of the INFORMATION STRUCTURE GROUND token of red in Figure 32 is 6216.26. These maximum amplitude values are plotted as the Y and X coordinates respectively of the point indicated by the “x” in the scatterplot in Figure 33. Maximum amplitudes of other INFORMATION STRUCTURE FOCUS versus INFORMATION STRUCTURE GROUND matched pairs are plotted similarly, and constitute the rest of the points shown in the scatterplot.
FIGURE 31. Waveform for the red mop pushes the white box with amplitude plot for red showing maximum amplitude.
FIGURE 32. Waveform for *the red shoe hits a blue teapot* with amplitude plot for *red* showing maximum amplitude
The scatter plot in Figure 33 and the Wilcoxon Signed-Rank Test show that the results for INFORMATION STRUCTURE FOCUS (IS-focus) versus INFORMATION STRUCTURE GROUND Max Amplitude 

Test Mean (ISfocus - ISground) =value

Hypothesized Value 0
Actual Estimate 956.515

Wilcoxon Signed-Rank Test

Test Statistic 90.000
Prob > |t| 0.038
Prob > t 0.019
Prob < t 0.981
(IS-ground) maximum amplitude are similar to those for mean amplitude. In particular, INFORMATION STRUCTURE FOCUS is greater than INFORMATION STRUCTURE GROUND at a statistically significant level for both mean and maximum amplitude. The significance level of $p<0.019$ for the maximum amplitude is not as dramatic as the $p<0.006$ level for the mean amplitude, but it is still significant and reinforces the same conclusions as the mean amplitude results. INFORMATION STRUCTURE FOCUS items have greater amplitude than their INFORMATION STRUCTURE GROUND counterparts, regardless of which measure of amplitude is used.

The example matched pair for DISCOURSE NEW versus DISCOURSE OLD is shown in Figure 34 and Figure 35. Maximum amplitude of 4435.88 for the DISCOURSE NEW token of green and maximum amplitude of 2519.64 for the DISCOURSE OLD token of green are indicated on the respective amplitude plots. These values appear again in the scatterplot in Figure 36 as the “x”. Maximum amplitude values of other DISCOURSE NEW versus DISCOURSE OLD matched pairs are also appear in the Figure 36 scatterplot with the maximum amplitude of the DISCOURSE NEW token of each pair as the Y-coordinate of a point and the maximum amplitude of its DISCOURSE OLD counterpart as the X-coordinate.
FIGURE 34. Waveform for a green lock crosses the divider with amplitude plot for green showing maximum amplitude
FIGURE 35. Waveform for the paperclip hits the green anchor with amplitude plot for green showing maximum amplitude.
FIGURE 36. Matched Pairs: DISCOURSE NEW vs OLD Max Amplitude

Maximum Amplitude
DISCOURSE NEW by DISCOURSE OLD

Test Mean (Dnew - Dold) =value
Hypothesized Value 0
Actual Estimate 726.139

Wilcoxon Signed-Rank Test
Test Statistic 150.000
Prob > |t| 0.125
Prob > t 0.063
Prob < t 0.937

In the Figure 36 scatterplot, the distribution of points is fairly even around the x=y line. While there are slightly more points above the x=y line than below it, and there are a few points for which the INFORMATION STRUCTURE FOCUS maximum amplitude is substan-
tially larger than the INFORMATION STRUCTURE GROUND maximum amplitude, there is not a clear trend. This is confirmed by the results of the Wilcoxon Signed-Rank test, which is not significant for either one or two tailed tests if 0.05 is taken as the cutoff for significance. It is interesting to note however, that the one tailed test, which if significant would indicate that DISCOURSE NEW items have greater amplitude than DISCOURSE OLD items, is much closer to being significant for maximum amplitude (p<0.063) than it was for either mean amplitude (p<0.219) or for duration (p<0.151). In summary, unlike the INFORMATION STRUCTURE FOCUS versus INFORMATION STRUCTURE GROUND distinction, the DISCOURSE NEW versus DISCOURSE OLD distinction does not have a significant effect on amplitude.

And finally, consider RELEVANT NON-UNIQUENESS versus not RELEVANT NON-UNIQUENESS. Waveform displays and accompanying amplitude plots indicating maximum amplitude are shown in Figure 37 and Figure 38 for the RELEVANTLY NON-UNIQUE versus not RELEVANTLY NON-UNIQUE example matched pair of the item white. The maximum amplitude of the RELEVANTLY NON-UNIQUE token of white Figure 37 is 2041.38. The maximum amplitude of the not RELEVANTLY NON-UNIQUE member of the matched pair in Figure 38 is 2371.59. The maximum amplitude values of this example matched pair are plotted in Figure 39 as the “x”. In the usual fashion, this example and the other RELEVANTLY NON-UNIQUE versus not RELEVANTLY NON-UNIQUE matched pairs are plotted in the Figure 39 scatterplot with the maximum amplitude of the RELEVANTLY NON-UNIQUE member of each pair as the Y-coordinate of a point and the maximum amplitude of the not RELEVANTLY NON-UNIQUE member of the pair as its X-coordinate.
FIGURE 37. Waveform for the red mop pushes the white box with amplitude plot for white showing maximum amplitude
FIGURE 38. Waveform for the blue oval hits the white bed with amplitude plot for white showing maximum amplitude
Visual inspection of the Figure 39 scatterplot would indicate, if anything, a counter-intuitive trend toward not RELEVANTLY NON-UNIQUE (RNU) item having greater maximum amplitude than RELEVANTLY NON-UNIQUE (not RNU) ones. However, the results of the
Wilcoxon Signed-Rank test show that there is no significant difference between RELEVANTLY NON-UNIQUE and not RELEVANTLY NON-UNIQUE items with respect to maximum amplitude. Note that this is completely compatible with what was found for the RELEVANT NON-UNIQUENESS versus not RELEVANT NON-UNIQUENESS distinction with respect to mean amplitude. The conclusion for RELEVANT NON-UNIQUENESS versus not RELEVANT NON-UNIQUENESS is that it has no effect on amplitude regardless of which measure is used.

The overall conclusion for amplitude is that INFORMATION STRUCTURE FOCUS has greater amplitude than INFORMATION STRUCTURE GROUND but that INFORMATION STRUCTURE status is the only factor that has an effect. This conclusion is supported by the results for both mean amplitude and maximum amplitude. The results of this section are summarized in Table 6.

**TABLE 6. Summary of Effect of Discourse Factors on Amplitude**

<table>
<thead>
<tr>
<th></th>
<th>Mean Amplitude (significance level)</th>
<th>Maximum Amplitude (significance level)</th>
</tr>
</thead>
<tbody>
<tr>
<td>INFORMATION STRUCTURE FOCUS</td>
<td>&gt;IS-ground (0.006)</td>
<td>&gt;IS-ground (0.025)</td>
</tr>
<tr>
<td>INFORMATION STRUCTURE - GROUND</td>
<td>&lt;IS-focus (0.006)</td>
<td>&lt;IS-focus (0.025)</td>
</tr>
<tr>
<td>DISCOURSE NEW</td>
<td>no effect (0.219)</td>
<td>no effect (0.063)</td>
</tr>
<tr>
<td>DISCOURSE OLD</td>
<td>no effect (0.219)</td>
<td>no effect (0.063)</td>
</tr>
</tbody>
</table>
5.1.3 Amplitude Integrals

The motivation behind investigating the effects of INFORMATION STRUCTURE status, DISCOURSE NEWNESS or OLDNESS and RELEVANT NON-UNIQUENESS with the integral of the amplitude of a word over its duration is to capture effects that might be masked by looking at duration or amplitude alone. If rather than attending to duration or amplitude of linguistic items independently, hearers attend to an overall level of energy in perceiving prominence, sufficient energy to be prominent could be achieved by duration or amplitude or various combinations. The amplitude integral should accommodate this sort of flexibility in producing prominence. The amplitude integrals were calculated using the trapezoidal method. Although there are many more sophisticated numerical methods for calculating integrals, there does not seem to be any obvious justification for using any of them when the amplitude function is unknown and only sampled points are available for calculation. Since the end goal is simply to have a number that represents the effects of both amplitude and duration in a way that would allow them to “trade-off,” the trapezoidal method seems as likely to yield reasonable results for this data as any other approach.

<table>
<thead>
<tr>
<th></th>
<th>Mean Amplitude (significance level)</th>
<th>Maximum Amplitude (significance level)</th>
</tr>
</thead>
<tbody>
<tr>
<td>RELEVANT NON-UNIQUENESS</td>
<td>no effect (0.207)</td>
<td>no effect (0.103)</td>
</tr>
<tr>
<td>not RELEVANT NON-UNIQUE-NESS</td>
<td>no effect (0.207)</td>
<td>no effect (0.103)</td>
</tr>
</tbody>
</table>
The amplitude integrals for the INFORMATION STRUCTURE FOCUS versus INFORMATION STRUCTURE GROUND example are depicted in Figure 40 and Figure 41 by the shaded area of the amplitude plots for the two tokens of the lexical item *red*. Both tokens of *red* are DISCOURSE OLD and RELEVANTLY NON-UNIQUE. The INFORMATION STRUCTURE FOCUS *red* in Figure 40 has an amplitude integral value of 827.941. The INFORMATION STRUCTURE GROUND member of this matched pair shown in Figure 41 has an amplitude integral value of 708.777. The point formed by using the amplitude integral value of 827.941 from the INFORMATION STRUCTURE FOCUS member of the pair as the Y-coordinate and the amplitude integral value of 708.777 from the INFORMATION STRUCTURE GROUND member as the X-coordinate is represented on the scatterplot in Figure 42, as before, as an “x”.

The other INFORMATION STRUCTURE FOCUS versus INFORMATION STRUCTURE GROUND matched pairs are plotted similarly with the amplitude integral of the INFORMATION STRUCTURE FOCUS member of the pair plotted on the Y-axis and the amplitude integral of its INFORMATION STRUCTURE GROUND partner plotted on the X-axis.
FIGURE 40. Waveform for the red mop pushes the white box with amplitude plot for red showing amplitude integral.
FIGURE 41. Waveform for the red shoe hits a blue teapot with amplitude plot for red showing amplitude integral.
The scatterplot of INFORMATION STRUCTURE FOCUS (IS-focus) versus INFORMATION STRUCTURE GROUND (IS-ground) in Figure 42 shows that the amplitude integrals for
INFORMATION STRUCTURE FOCUS items are significantly larger than for their INFORMATION STRUCTURE GROUND counterparts. This is not especially surprising, since the effect of being INFORMATION STRUCTURE FOCUS versus INFORMATION STRUCTURE GROUND was strong for both duration and amplitude, and in both cases the relevant number for INFORMATION STRUCTURE FOCUS items was greater than for their INFORMATION STRUCTURE GROUND counterparts. In fact, the Wilcoxon Signed-Rank Test shows that the difference is even more significant for the amplitude integral (p<0.002) than for either duration (p<0.003) or amplitude (p<0.006 mean, p<0.025 max) alone.

In Figure 43 and Figure 44 the two tokens of *green* that form the example matched pair for DISCOURSE NEW versus DISCOURSE OLD are displayed with amplitude plots that show the amplitude integral as the shaded area under the amplitude function. As in previous uses of this example, both tokens of *green* are INFORMATION STRUCTURE FOCUS and not RELATIVELY NON-UNIQUE. The DISCOURSE NEW token of *green* in Figure 43 has an amplitude integral value of 929.329, while the DISCOURSE OLD token of *green* in Figure 44 has an amplitude integral value of 464.634. These amplitude integral values, 929.329 and 464.634, appear in the Figure 45 scatterplot in the usual manner as the Y and X-coordinates respectively of the point indicated by the “x”.
FIGURE 43. Waveform for a green lock crosses the divider with amplitude plot for green showing amplitude integral.
FIGURE 44. Waveform for the paperclip hits the green anchor with amplitude plot for green showing amplitude integral.
One can see in Figure 45 that the scatterplot for amplitude integrals might show a slight trend toward DISCOURSE NEW items having greater amplitude integrals than DISCOURSE OLD.
OLD items. However, if there is such a trend it is sufficiently weak such that the results for the Wilcoxon Signed-Rank Test are not significant for either one or two tailed versions of the test. The one tailed test that would be interpreted as showing that DISCOURSE NEW items have larger amplitude integrals than DISCOURSE OLD items is significant at the p<0.068 level -- too big for significance given a 0.05 cutoff, yet is more significant than the p<0.219 level of significance for the mean amplitude or the p<0.151 level of significance for duration. This is similar to the result for the comparison of DISCOURSE NEW versus DISCOURSE OLD maximum amplitude.

Figure 46 and Figure 47 show the tokens of white that constitute the example matched pair for RELEVANT NON-UNIQUENESS versus not RELEVANT NON-UNIQUENESS with waveform displays and accompanying amplitude plots appropriately shaded to represent the amplitude integrals. The RELEVANTLY NON-UNIQUE member of the pair in Figure 46 has an amplitude integral value of 332.405, while the amplitude integral value of its not RELEVANTLY NON-UNIQUE counterpart in Figure 47 is 300.579. In the Figure 48 scatterplot the example matched pair is represented by the “x”. The amplitude integral value of 332.405 for the RELEVANTLY NON-UNIQUE instance of white is plotted as the Y-coordinate of this point and the amplitude integral value, 300.579 is plotted as its X-coordinate. The other RELEVANTLY NON-UNIQUE versus not RELEVANTLY NON-UNIQUE matched pairs are represented on the Figure 48 scatterplot in a similar fashion with the amplitude integral of the RELEVANTLY NON-UNIQUE member of each pair as the Y-coordinate of a point and the amplitude integral of the not RELEVANTLY NON-UNIQUE member as the point’s X-coordinate.
FIGURE 46. Waveform for the red mop pushes the white box with amplitude plot for white showing amplitude integral.
FIGURE 47. Waveform for the blue oval hits the white bed with amplitude plot for white showing amplitude integral
FIGURE 48. Matched Pairs: RELATIVELY NON-UNIQUE vs NOT Amplitude Integral

**Test Mean (RNU - not RNU) = value**

Hypothesized Value 0
Actual Estimate -109.89

**Wilcoxon Signed-Rank Test**

Test Statistic -15.000
Prob > |t| 0.206
Prob > t 0.897
Prob < t 0.103

RELEVANT NON-UNIQUENESS (RNU), which surprisingly showed no effect for either duration or amplitude, also shows no effect for amplitude integrals, as can be seen in
Figure 48. The points on the scatterplot would if anything indicate the counter-intuitive situation of the not RELEVANTLY NON-UNIQUE (not RNU) items having larger amplitude integrals than RELEVANTLY NON-UNIQUE items. However, neither the counter-intuitive trend suggested by the scatterplot nor the one that might have been expected is significant based on the results of the Wilcoxon Signed-Rank Test. The conclusion for the RELEVANT NON-UNIQUENESS versus not RELEVANT NON-UNIQUENESS distinction is that it has no effect on overall energy as represented by the amplitude integral.

The conclusion for the amplitude integral is that only INFORMATION STRUCTURE has a significant effect. INFORMATION STRUCTURE FOCUS items have greater overall energy than INFORMATION STRUCTURE GROUND items. Being DISCOURSE NEW versus DISCOURSE OLD or being RELEVANTLY NON-UNIQUE versus not RELEVANTLY NON-UNIQUE has no effect. The summary results for amplitude integral is shown in Table 7.

TABLE 7. Summary of Effect of Discourse Factors on Amplitude Integral

<table>
<thead>
<tr>
<th></th>
<th>Amplitude Integral (significance level)</th>
</tr>
</thead>
<tbody>
<tr>
<td>INFORMATION STRUCTURE FOCUS</td>
<td>&gt;IS-ground (0.002)</td>
</tr>
<tr>
<td>INFORMATION STRUCTURE GROUND</td>
<td>&lt;IS-focus (0.002)</td>
</tr>
<tr>
<td>DISCOURSE NEW</td>
<td>no effect (0.068)</td>
</tr>
<tr>
<td>DISCOURSE OLD</td>
<td>no effect (0.068)</td>
</tr>
</tbody>
</table>
5.1.4 Summary

INFORMATION STRUCTURE FOCUS is longer and louder than INFORMATION STRUCTURE GROUND. This is supported by the results of the investigation of duration, amplitude and amplitude integral. RELEVANT NON-UNIQUENESS, somewhat surprisingly, has no effect on either amplitude or duration alone and no effect on the amplitude integral. So it appears from this data that RELEVANTLY NON-UNIQUE items are not consistently louder or longer than their not RELEVANTLY NON-UNIQUE counterparts, and they also do not appear to have more overall energy.

<table>
<thead>
<tr>
<th></th>
<th>Amplitude Integral (significance level)</th>
</tr>
</thead>
<tbody>
<tr>
<td>RELEVANT NON-UNIQUENESS</td>
<td>no effect</td>
</tr>
<tr>
<td></td>
<td>(0.103)</td>
</tr>
<tr>
<td>not RELEVANT NON-UNIQUENESS</td>
<td>no effect</td>
</tr>
<tr>
<td></td>
<td>(0.103)</td>
</tr>
</tbody>
</table>

**TABLE 8. Summary of Effects of Discourse Factors on Duration, Amplitude and “Energy”**

<table>
<thead>
<tr>
<th></th>
<th>Duration</th>
<th>Amplitude</th>
<th>“Energy” (Amplitude Integral)</th>
</tr>
</thead>
<tbody>
<tr>
<td>INFORMATION STRUCTURE FOCUS</td>
<td>&gt;IS-ground</td>
<td>&gt;IS-ground</td>
<td>&gt;IS-ground</td>
</tr>
<tr>
<td>INFORMATION STRUCTURE GROUND</td>
<td>&lt;IS-focus</td>
<td>&lt;IS-focus</td>
<td>&lt;IS-focus</td>
</tr>
</tbody>
</table>
In conclusion, the analysis in this section has demonstrated that when effects of the three focus interpretations, INFORMATION STRUCTURE FOCUS, DISCOURSE NEW, and RELEVANT NON-UNIQUENESS, are each isolated by controlling for the other two, only differences in INFORMATION STRUCTURE status reliably correlate with differences in duration and amplitude. INFORMATION STRUCTURE FOCUS items are longer and louder than INFORMATION STRUCTURE GROUND items while there is no effect for either DISCOURSE NEW/OLD or RELEVANT NON-UNIQUENESS.

5.2 Duration and Amplitude within INFORMATION STRUCTURE FOCUS Noun Phrase

Given the significant effects of INFORMATION STRUCTURE status on duration and amplitude, one might wonder how these effects distribute within a focussed phrase. Particularly since the distribution of focus effects can provide evidence for or against focus projection

<table>
<thead>
<tr>
<th></th>
<th>Duration</th>
<th>Amplitude</th>
<th>“Energy” (Amplitude Integral)</th>
</tr>
</thead>
<tbody>
<tr>
<td>DISCOURSE NEW</td>
<td>no effect</td>
<td>no effect</td>
<td>no effect</td>
</tr>
<tr>
<td>DISCOURSE OLD</td>
<td>no effect</td>
<td>no effect</td>
<td>no effect</td>
</tr>
<tr>
<td>RELEVANT NON-UNIQUENESS</td>
<td>no effect</td>
<td>no effect</td>
<td>no effect</td>
</tr>
<tr>
<td>not RELEVANT NON-UNIQUENESS</td>
<td>no effect</td>
<td>no effect</td>
<td>no effect</td>
</tr>
</tbody>
</table>

TABLE 8. Summary of Effects of Discourse Factors on Duration, Amplitude and “Energy”
accounts of focus marking in English. In the investigation reported in Section 5.1, no distinction was made among the lexical items that make up an INFORMATION STRUCTURE FOCUS or INFORMATION STRUCTURE GROUND phrase. Even though the matched pairs were matched pairs of words, the coding strategy was largely phrase based. Every lexical item that was part of an INFORMATION STRUCTURE FOCUS phrase was coded as INFORMATION STRUCTURE FOCUS and any lexical item contained in an INFORMATION STRUCTURE GROUND phrase was marked as INFORMATION STRUCTURE GROUND. This means that, for example, for an INFORMATION STRUCTURE FOCUS NP the noun, the adjective and the determiner are coded as INFORMATION STRUCTURE FOCUS and were treated the same in the matched pair analysis.

Focus projection accounts would predict, however, that the components of a focussed phrase would not be uniformly marked. These phonological accounts of focus marking are concerned with what size and nature of constituent a single pitch accent can mark as focussed (e.g. Selkirk (1984) Gussenhoven (1983)). The core of the problem is that items with broad and narrow focus interpretations look very similar phonologically, especially with respect to pitch accent placement. To address this problem both Gussenhoven (1983) and Selkirk (1984) have focus projection or focus domain formation rules that permit a pitch accent on a noun in an NP direct object to ambiguously mark the noun itself, the NP, the VP and the whole sentence as focus. Given this type of account, in so far as increased duration and amplitude can be taken to be indicative of pitch accenting, one might expect that only the noun of a focussed NP or direct object within a focussed VP would show greater duration and amplitude as compared to non-focussed lexical items.
Other lexical items in the focussed constituents would not be expected to show any increased duration or amplitude.

The Animation Corpus task was not designed to elicit VP foci but NP foci are common in the corpus. This section investigates the question of whether it is only the noun in an INFORMATION STRUCTURE FOCUS NP that has greater duration and amplitude compared to INFORMATION STRUCTURE GROUND items. Only INFORMATION STRUCTURE will be considered since it is the only one of the focus interpretations tested that had significant duration and amplitude effects. The same matched pairs are used as for the previous duration and amplitude analysis, but they are separated by part of speech, and only NP focus cases are considered (i.e. adjective focus cases are excluded).

The number of nouns is small (N=6) due to the control on newness built into the Animation Corpus task. As was discussed in Chapter 4, in order to insure that DISCOURSE NEW characters were really new and not remembered from previous animations, shapes were only repeated every 10 animations so the subjects simply did not have the opportunity to produce very many tokens of the same word. In addition, the strict matching of pairs of words on multiple criteria reduces the number of usable items. However, the relatively small number of noun pairs is not as serious a problem as it might seem for two reasons. First, the Wilcoxon Signed Rank test in theory can be valid with an N as small as 6, and presumably if the pattern is clear enough will still show significance appropriately even under this condition. Second, the outcome for the nouns is the least interesting. We expect there to be an effect for the nouns; the real question is whether there is an effect for anything else. The results just for the adjectives and determiners (for which there are...
many more tokens, N=19 for adjectives and N=23 for determiners) would be sufficient to answer the question, even if the results for the nouns were taken to be inconclusive.

Table 9 shows the results of the analysis separating the three parts of speech. The results are significant for two of the four measures for nouns. Duration of INFORMATION STRUCTURE FOCUS nouns is greater than that of INFORMATION STRUCTURE GROUND nouns at a significance level of p<.016. Maximum amplitude of INFORMATION STRUCTURE FOCUS nouns is also significantly greater than for INFORMATION STRUCTURE GROUND nouns (p<.031). For adjectives and determiners, the INFORMATION STRUCTURE FOCUS items are greater on all four measures than their INFORMATION STRUCTURE GROUND counterparts at a significance level better than p<.05. From these statistically significant differences, particularly for the adjectives and determiners, it seems that duration and amplitude effects extend to all the components of a focussed NP and are not restricted to the noun. This suggests some reevaluation of the focus projection account may be required.

<table>
<thead>
<tr>
<th></th>
<th>Duration</th>
<th>Mean Amplitude</th>
<th>Maximum Amplitude</th>
<th>“Energy” (Amplitude Integral)</th>
</tr>
</thead>
<tbody>
<tr>
<td>Nouns</td>
<td>ISf &gt; ISg (p&lt;0.016)</td>
<td>Not Significant (p&lt;0.344)</td>
<td>ISf &gt; ISg (p&lt;0.031)</td>
<td>Not Significant (p&lt;0.078)</td>
</tr>
<tr>
<td>Adjectives</td>
<td>ISf &gt; ISg (p&lt;0.044)</td>
<td>ISf &gt; ISg (p&lt;0.038)</td>
<td>ISf &gt; ISg (p&lt;0.045)</td>
<td>ISf &gt; ISg (p&lt;0.016)</td>
</tr>
<tr>
<td>Determiners</td>
<td>ISf &gt; ISg (p&lt;0.032)</td>
<td>ISf &gt; ISg (p&lt;0.012)</td>
<td>ISf &gt; ISg (p&lt;0.037)</td>
<td>ISf &gt; ISg (p&lt;0.025)</td>
</tr>
</tbody>
</table>
There are several ways to think about these results in relation to the ambiguous focus/focus projection story. One could decide that duration and amplitude are not necessarily correlates of pitch accents. The question then becomes what phonological entity duration and amplitude do correlate with. One possibility is to accept Ladd’s (1996) argument that the primary expression of focus in English is metrical rather than accentual. If that were the case the duration and amplitude effects observed in the English Animation Corpus could reasonably be analyzed as cues to metrical structure and the relative prominence of units within the metrical structure rather than as correlates of pitch accent. Such an account would have to posit association of increased duration and amplitude with a strong node in the metrical structure dominating the focussed constituent. Under this view INFORMATION STRUCTURE FOCUS might be marked ambiguously by pitch accents and non-ambiguously by duration and amplitude, or we could decide that pitch accents do not mark INFORMATION STRUCTURE FOCUS at all, and in fact serve to mark something else while duration and amplitude as correlates of metrical strength mark INFORMATION STRUCTURE status.

Another possibility is that a different notion of pitch accent could be adopted in which the pitch accent can be seen as aligning with constituents larger than a word, with the span being phonetically implemented by duration and amplitude, while the phonetic implementation of the F0 component could still be expected to align an F0 event with a particular syllable of a particular word. Or perhaps the duration and amplitude effects could be modeled as phonological spreading bounded by the focussed constituent. Regardless of what solution is adopted to account for these duration and amplitude effects, the idea that focus
marking is ambiguous between smaller and larger foci and the idea that focus marking is localized on a single syllable of a single word have been put in question by the results presented here.