A Computational Phonology of Russian

by
Peter A. Chew

ISBN: 1-58112-178-4

DISSERTATION.COM

Parkland, FL • USA • 2003

A Computational Phonology of Russian

Dissertation.com
USA • 2003

ISBN: 1-58112-178-4
www.Dissertation.com/library/1121784a.htm

A Computational Phonology of Russian

Peter Chew
Jesus College, University of Oxford
D. Phil. dissertation, Michaelmas 1999

Abstract

This dissertation provides a coherent, synchronic, broad-coverage, generative phonology of Russian. I test the grammar empirically in a number of ways to determine its goodness of fit to Russian. In taking this approach, I aim to avoid making untested (or even incoherent) generalizations based on only a handful of examples. In most cases, the tests show that there are exceptions to the theory, but at least we know what the exceptions are, a baseline is set against which future theories can be measured, and in most cases the percentage of exceptional cases is reduced to below 5%.

The principal theoretical outcomes of the work are as follows. First, I show that all of the phonological or morphophonological processes reviewed can be described by a grammar no more powerful than context-free.

Secondly, I exploit probabilistic constraints in the syllable structure grammar to explain why constraints on word-marginal onsets and codas are weaker than on word-internal onsets and codas. I argue that features such as [±initial] and [±final], and extraprosodicity, are unnecessary for this purpose.

Third, I claim that /v/ should be lexically unspecified for the feature [±sonorant], and that the syllable structure grammar should fill in the relevant specification based on its distribution. This allows a neat explanation of the voicing assimilation properties of /v/, driven by phonotactics.

Fourth, I argue that jers in Russian should be regarded as morphological objects, not segments in the phonological inventory. Testing the grammar suggests that while epenthesis cannot be regarded as a major factor in explaining vowel-zero alternations, it might be used to explain a significant minority of cases.

Fifth, I suggest that stress assignment in Russian is essentially context-free, resulting from the intersection of morphological and syllable structure constraints. I show that my account of stress assignment is simpler than, but just as general as, the best of the three existing theories tested.

Finally, this dissertation provides new insight into the nature and structure of the Russian morphological lexicon. An appendix of 1,094 morphemes and 1,509 allomorphs is provided, with accentual and jer-related morphological information systematically included.

A Computational Phonology of Russian

by

Peter Chew

University of Oxford
Jesus College
Michaelmas 1999

Thesis submitted for the degree of Doctor of Philosophy
at the University of Oxford

Acknowledgements

I would like to thank my supervisor, John Coleman, for his help. Without his encouragement and support even before I embarked upon this research, I would doubtless now be a well-paid but bored chartered accountant. Auditing linguistic theories has proved to be more rewarding in many ways than auditing financial statements, and I am confident that the choice of leaving my previous job to pursue this research was the right one.

It would not have been possible to complete this D. Phil. without the support of my wife, Lynn. She has always been there to give practical suggestions, as a sounding board for ideas, and simply as a partner in life, sharing encouraging and discouraging times together. God could not have given me a better wife.

My parents have also been a great practical help, babysitting almost weekly, having us round for meals, and generally helping reduce the stress in our lives. Although Jonathan, who was born 15 months before I submitted this thesis, has taken time from my studies, we are very grateful for his arrival. I cannot think of a better way of spending my time, and I cannot imagine a better son.

A number of people have read drafts of my work or listened to me, giving helpful advice which enabled me to sharpen my thoughts and improve the way in which I expressed them. Thanks (in alphabetical order) to Dunstan Brown, Bob Carpenter, Andrew Hippisley, Mary MacRobert, Stephen Parkinson, Burton Rosner, Irina Sekerina, Andrew Slater, and Ian Watson. Andrew Slater has also provided invaluable technical support. I often feel that he puts the rest of us to shame with his good humour, helpfulness, and a constant willingness to go the extra mile.

My friends at the Cherwell Vineyard Christian Fellowship have provided a dependable support network which has kept Lynn and me going through not always easy times. First and foremost, they have encouraged us to keep looking towards the one without whom we can do nothing. However, I know I will also look back on the laughs we have had with Richard and Janet Remmington, Evan and Eowyn Robertson, Judy Irving, and others, on Thursday evenings with fond memories.

Finally, I would like to thank my college, Jesus College, for providing very generous financial support throughout my time at Oxford. And without the financial support of the Arts and Humanities Research Board (formerly the British Academy), I would not have undertaken this research project in the first place.

List of abbreviations and symbols

General symbols

//	enclose phonemic representations, e.g. /sdat^j/
[]	enclose phonetic representations, e.g. [zdat^j]
/s$_c$+misl$_{rn}$+$_{in}$/	denotes morphological tokenization; subscripts classify individual morphs
+	morpheme boundary
.	syllable boundary
ˈ	denotes word-stress in IPA transcriptions (stress on the vowel to the right)
/misl/$_{rn}$	denotes a single morpheme (classificatory subscript is outside obliques)
σ	syllable
∅	the empty string
anter	anterior
C	any consonant
CFG	context-free grammar
cons	consonantal
cont	continuant
coron	coronal
DCG	(Prolog) Definite Clause Grammar
del_rel	delayed release
init	initial
later	lateral
OT	Optimality Theory
PSG	phrase structure grammar
sonor	sonorant
SSG	Sonority Sequencing Generalization
V	any vowel
vfv	vocal fold vibration
voc	vocalic

Symbols used in morphological tokenization

r*	root	d	durative process
s	suffix	r*	resultative process
c*	clitic	i	iterative process
i	inflectional ending	c*	completed process
p	pronominal		
a	adjectival		
n	substantival		
v	verbal		

*No ambiguity arises with respect to the use of non-unique symbols, because the meaning of each symbol is also dependent on its position; full details are given in section 3.2.1.2.

Table of contents

Table of figures

List of tables

Chapter 1: Introduction

1.1 Introduction

This dissertation provides a coherent, synchronic, broad-coverage, generative account of Russian phonology. By 'broad-coverage', I mean that it will cover a number of phonological phenomena (stress assignment, syllabification, vowel-zero alternations, word-final devoicing, voicing assimilation, vowel reduction, and consonant-vowel interdependencies) within a single constrained grammar. While I have not attempted to deal exhaustively with all the phonological problems of interest in Russian (for example, I do not attempt to account for all morphophonological alternations), the current work covers those areas which have attracted the most attention in the literature on Russian phonology.

While all these aspects of Russian phonology have been richly documented, generally they have been dealt with in isolation; the one notable exception to this is Halle's (1959) *Sound Pattern of Russian*. The following quotation (op. cit., p. 44) serves to show that Halle's account of Russian phonology is also intended to be broad-coverage in the sense just outlined:

> When a phonological analysis is presented, the question always arises as to what extent the proposed analysis covers the pertinent data. It is clearly impossible in a description to account for all phonological manifestations in the speech of even a single speaker, since the latter may (and commonly does) use features that are characteristic of different dialects and even foreign languages. (E.g., a speaker of Russian may distinguish between nasalized and nonnasalized vowels in certain [French] phrases which form an integral part of his habitual conversational repertoire.) If such facts were to be included, all hopes for a systematic description would have to be abandoned. It is, therefore, better to regard such instances as deviations to be treated in a separate section and to restrict the main body of the grammar to those manifestations which can be systematically described.

The aim of the current work is thus substantially the same as that of Halle (1959). However, in the forty years since then there have been a number of advances, both linguistic and technological, which allow us to take a fresh (and perhaps more rigorous) look at some of the same phenomena which Halle and others attempted to describe. In the late 1950s and early 1960s Chomsky and co-workers pioneered work in developing a formal theory of language (Chomsky 1959, 1963, 1965); this work established clearly-defined links between linguistics, logic and mathematics, and was also foundational in computer science in the sense that the principles it established have also been applied in understanding computer programming languages. These advances make it possible to formulate a theory of Russian phonology, just as Halle did, but to test it empirically by implementing the theory as a computer program and using it to process very large numbers of words. Moreover, since the technological advances which make it possible to do this owe a great deal to Chomsky's work, the transition from generative grammar to computational grammar can be a comparatively straightforward one.

One of the defining features of generative grammar is the emphasis on searching for cross-linguistic patterns. Without denying the value of language-specific grammar, Chomsky and Halle (1968) (to many the canonical work of generative phonology) illustrates this thinking:

> ...we are not, in this work, concerned exclusively or even primarily with the facts of English as such. We are interested in these facts for the light they shed on linguistic theory (on what, in an earlier period, would have been called "universal grammar") and for what they suggest about the nature of mental processes in general… We intend no value judgment here; we are not asserting that one *should* be primarily concerned with universal grammar and take an interest in the particular grammar of English only insofar as it provides insight into universal grammar and psychological theory. We merely want to make it clear that this is our point of departure in the present work; these are the considerations that have determined our choice of topics and the relative importance given to various phenomena. (p. viii)

The emphasis on cross-linguistic generalization, characteristic of Chomsky's work, has characterized generative linguistics ever since: indeed, there is a considerable branch of linguistics (Zwicky 1992 is an example) which abstracts completely away from language-specific data. (This branch deals in what Zwicky 1992: 328 refers to as 'frameworks' as opposed to 'theories'.) While frameworks have their place (indeed, a theory cannot exist without a framework), the difficulty is always that frameworks cannot be verified without theories. In this light, Chomsky and Halle (1968) claimed to establish both a cross-linguistic framework and a theory about English phonology.

The focus of this description is on ensuring that the phonology of Russian proposed is both internally consistent and descriptively adequate — that is, that it makes empirically correct predictions about Russian — rather than on attempting to develop any particular linguistic framework. Exciting possibilities are open in this line of research thanks to the existence of computer technology. It is possible to state grammatical rules in a form which has the rigour required of a computer program, and once a program is in place, large corpora can be quickly processed. Thus the phonology of Russian presented here is 'computational' simply because of the advantages in speed and coverage that this approach presents.

Establishing that a linguistic theory can be implemented as a computer program and verifying its internal consistency in this way is a valuable exercise in itself, but non-computational linguists may be sceptical: some may argue that this kind of approach does not contribute anything to linguistics per se. Whether or not this is criticism is well-founded (and I believe it is not), I hope that this dissertation will satisfy even the more stringent critics by making a number of key contributions to linguistic knowledge. These are as follows.

First, I propose that both the distribution of /v/ and its behaviour with respect to voicing assimilation can be explained if /v/, unlike all other segments in the phonological inventory of Russian, is lexically unspecified for the feature [±sonorant]. The syllable structure rules determine whether /v/ is [+sonorant] or [−sonorant], and this in turn determines how /v/ assimilates in voice to adjacent segments.

Second, I suggest that the greater latitude allowed in word-marginal onsets and codas, which is a feature of Russian and other languages (cf. Rubach and Booij 1990), can be explained naturally by a probabilistic syllable structure grammar. This approach allows features such as [±initial] and [±final] (cf. Dirksen 1993) to be dispensed with.

Third, I show that vowel-zero alternations in Russian cannot fully be explained by a Lexical-Phonology-style account (such as that proposed by Pesetsky ms 1979) alone, nor can they be the result of epenthesis alone. I show empirically that a combination of factors, including (1) the morphophonological principles discovered by Pesetsky, (2) epenthesis, and (3) etymology, governs vowel-zero alternations.

Fourth, I show that Russian stress can be accounted for with a high rate of accuracy by existing generative theories such as that of Melvold (1989), but I suggest a simpler theory which accounts for the same data with as good a rate of accuracy. The theory which I propose regards stress assignment as resulting from the interaction of morphological and syllable structure: existing generative theories do not acknowledge syllable structure as playing any role in Russian stress assignment. An integral part of my theory is a comprehensive inventory of morphemes together with the accentual information which is lexically specified for each morpheme. The inventory which I propose, which is arrived at by computational inference, includes 1,094 morphemes and 1,509 allomorphs, while the longest existing list of this type, as far as I am aware, is the index of approximately 250 suffixes in Red'kin (1971).

The structure of this dissertation is as follows. In this chapter, I set out in detail the concepts which are foundational to the whole work: the role which computation plays in my work (1.2), the framework which I use (1.3), and the methodology which underlies my work (1.4). Then, I discuss in detail aspects of the syllable structure and morphological structure of Russian in Chapters 2 and 3 respectively, in each case developing a formally explicit grammar module which can be shown to be equivalent to a finite state grammar. Chapter 4 describes in detail three theories of stress assignment in Russian. These are tested computationally to ascertain which is the most promising. Each of Chapters 2-4 begins with a section reviewing the relevant literature. Finally, in Chapter 5, I describe how the principal features of the preferred theory from Chapter 4 can be incorporated into a synthesis of the grammars developed in Chapters 2 and 3. The result is an integrated, internally consistent, empirically well-grounded grammar, which accounts for a variety of different aspects of Russian phonology.

1.2 Why 'computational' linguistics?

In this dissertation, computation is used as a tool. Any tool has limitations, of course: a large building cannot be built with a power drill alone, and, to be sure, there are problems in linguistics which computation is ill-suited to solve. On the other hand, anyone who has a power drill will try to find appropriate uses for it. Likewise, I aim to use computation for the purposes for which it is best suited. This, then, is not a dissertation 'about' computational linguistics; it is a dissertation that uses computation as a tool in linguistics.

What, then, *are* the strengths of computational tools in linguistics? Shieber (1985: 190-193), noting that the usefulness of computers is often taken for granted by computational linguists, lists three roles that the computer can play in the evaluation of linguistic analyses: the roles of straitjacket (forcing rigorous consistency and explicitness, and clearly delineating the envelope of a theory), touchstone ('indicating the correctness and completeness of an analysis'), and mirror ('objectively reflecting everything in its purview'). In short, the process of implementing a grammar computationally forces one to understand in detail the mechanisms by which a grammar assigns structure. Shieber states, for example, that

> …we have found that among those who have actually attempted to write a computer-interpretable grammar, the experience has been invaluable in revealing real errors that had not been anticipated by the Gedanken-processing typically used by linguists to evaluate their grammars — errors usually due to unforeseen interactions of various rules or principles. (p. 192)

This has also been my experience in developing the current phonology of Russian. In particular, areas such as stress assignment involve the interaction of a number of different grammar modules, and, as Shieber states, 'decisions in one part of the grammar, while internally consistent, may not cohere with interacting decisions in another part' (Shieber 1985: 190). Problems of this kind cannot always feasibly be foreseen without actually implementing and testing a theory on a corpus of data.

Another perhaps self-evident strength of computers is their ability to process large volumes of data quickly: once a grammar has been implemented, the processing can take place without intensive effort on the part of the researcher. While in principle generative theories can be implemented and tested by hand, the volume of data that typically has to be processed to achieve significant results means that this is an extremely tedious and time-consuming, if not impracticable, task. Clearly, computational techniques shift the burden for the researcher from data processing to the more interesting task of developing theories, identifying exceptions quickly, and debugging the theory as appropriate.

Because the discipline of computational linguistics is still relatively young, it is perhaps understandable that many existing theories have neither been implemented nor tested computationally, but now that the means to validate theories are widely available, it is less justifiable for new theories still to be proposed in linguistics *without* being empirically tested: ‘the widespread practice of testing a few interesting cases is unreliable and is no substitute for an exhaustive check’ (Bird 1995: 14). It seems that at this stage in linguistic research, the efforts of linguists would be better directed towards implementing and testing existing theories rather than proposing new alternatives, since otherwise it cannot be demonstrated that the new alternatives measure up any better to the criteria of coverage, constrainedness and ability to integrate than the theories which they replace.

It is also worth noting the limitations of computational analysis (which I set as the limits for this dissertation). Ultimately, computers follow instructions rather than making judgements, and while they are very good at evaluating grammars for consistency and descriptive adequacy, they cannot test for explanatory adequacy unless the programmer supplies the necessary information (that is, a standard against which to measure the accuracy of structures assigned by a grammar to strings). The judgement about the nature of the correct structures is a question of psychology, and therefore I do not claim that the current phrase-structure context-free phonology of Russian is a psychological model. In this, my approach is exactly the same as that of Gazdar, Klein, Pullum and Sag (1985):

> We make no claims, naturally enough, that our grammatical theory is *eo ipso* a psychological theory. Our grammar of English is not a theory of how speakers think up things to say and put them into words. Our general linguistic theory is not a theory of how a child abstracts from the surrounding hubbub of linguistic and nonlinguistic noises enough evidence to gain a mental grasp of the structure of a natural language. Nor is it a biological theory of the structure of an as-yet-unidentified mental organ. It is irresponsible to claim otherwise for theories of this general sort…
>
> Thus we feel it is possible, and arguably proper, for a linguist (*qua* linguist) to ignore matters of psychology. But it is hardly possible for a psycholinguist to ignore language… If linguistics is truly a branch of psychology (or even biology), as is often unilaterally asserted by linguists, it is so far the branch with the greatest pretensions and the fewest reliable results… So far, linguistics has not fulfilled its own side of the interdisciplinary bargain. (p. 5)

1.3 The framework

1.3.1 Phrase-structure grammar

In this dissertation, phonology and morphology, as modules of grammar, have the function of enumerating or generating (the words of a) language. This view of grammatical modules is entirely in accordance with traditional generative linguistics (e.g. Chomsky and Miller 1963: 283-285). More precisely, a phonological grammar should be able to generate all and only the phonological words of a natural language; similarly, a word-formation grammar should enumerate all the morphological words (p-forms, in the terminology of Zwicky 1992: 334) of a natural language.[1] The same

---

[1] As noted by Booij and Rubach (1984), there may well not be a one-to-one mapping between 'morphological words' and 'phonological words' — well-known examples from Russian are preposition-noun phrases, all of which have a single stress (e.g. зá руку /zˈa ruku/ 'by the hand') and are thus considered to function as a single phonological word, but two morphological words.

grammar that enumerates the forms of a language should also be able to assign them a structural description (that is, parse them). These functions are clearly fulfilled by phrase-structure grammars (PSGs), since in a PSG each rule can equivalently be thought of as a partial structure, and each derivation can be represented as a directed graph.

The ability of a grammar to parse (that is, provide *some* structural description for the word) does not necessarily imply its ability to parse *correctly*. As Chomsky and Miller (1963: 297) state, 'we have no interest, ultimately, in grammars that generate a natural language correctly but fail to generate the correct set of structural descriptions'. A grammar which is able to assign structural descriptions to all relevant well-formed utterances in a language is said to meet the condition of *descriptive adequacy*, while a grammar which meets the more stringent requirement of assigning correct structural descriptions to all well-formed utterances is said to meet the condition of *explanatory adequacy*. In general, it is considerably harder to prove or disprove a grammar's explanatory adequacy than its descriptive adequacy, since the former is a matter not just of linguistic data but of psychology as well (Chomsky 1965: 18-27). Moreover, it is important to realize that a parse should not necessarily be considered incorrect just because it was unanticipated: such a parse may in fact be a possible but unlikely parse. These factors all mean that establishing whether a given grammar assigns correct structural descriptions is not always straightforward, and is often a matter of judgement.

Conversely, English words of the form 'non-*X*' (where *X* stands for an adjective) are a single morphological word, but two phonological words.

Essentially, there are three good reasons for formulating a theory within the framework of PSG. First, PSGs are the standard means of assigning hierarchical constituent structure to strings, which is widely and uncontroversially regarded as an important function of linguistics. The literature on phrase-structure grammar has been developed over approximately 40 years, and owes much to Chomsky's interest in establishing a formal foundation for linguistics (e.g. Chomsky 1959, Chomsky and Miller 1963, Chomsky 1963).

A second strength of the PSG formalism is that it has a straightforward declarative interpretation. Phrase-structure grammar 'rules' can equally validly be seen as 'partial descriptions of surface representations' or 'descriptions of information structures', in Brown, Corbett, Fraser, Hippisley and Timberlake's (1996) terminology. Specifically, context-free phrase-structure rules can be represented graphically as tree structures (Coleman 1998: 99).

Third, there is a transparent relationship between PSGs and Definite Clause Grammars (DCGs)[2]. This is perhaps the greatest advantage of using the PSG formalism, because it means that a PSG can easily be implemented and tested computationally. DCGs are a particular type of formalism available as part of the programming language Prolog. For details of the workings of Prolog and DCGs, the reader is invited to refer to a textbook on Prolog, such as Clocksin and Mellish (1981). Here, it is sufficient to appreciate that DCGs can fulfil the functions of parsing and generation, because Prolog is a declarative programming language. Thus, if a

[2] DCGs are capable of defining recursively enumerable languages and CFGs are capable of defining only context-free languages (which are a subset of the set of recursively enumerable languages). Thus, to be more precise, the type of DCG used to implement the theory proposed in this dissertation is a restricted type of DCG.

particular grammar is implemented as a DCG, it is possible to test the grammar computationally to determine whether it 'describe[s] all and only the possible forms of a language' (Bird, Coleman, Pierrehumbert and Scobbie 1992). Throughout this dissertation, I describe computational tests of this kind to determine whether the different aspects of the grammar are accurate representations of the facts of the language.

1.3.2 Context-free grammar

Having established in section 1.3.1 why I use the framework of PSG, I now move on to explain the significance of my claim that nothing more powerful than a context-free grammar (CFG) is necessary to describe the facts of Russian phonology. The claim that CFG is sufficient is in contrast to McCarthy (1982), for example, who claims that phonology is context-sensitive (p. 201). (Coleman 1998: 81 observes that his phonology is an unrestricted rewriting system, since it is a context-sensitive grammar with deletion: see McCarthy's (1) on p. 201.) In other respects, however, McCarthy's aim is comparable to mine: McCarthy aims to provide 'a fair degree of coverage, particularly in Hebrew phonology and Arabic morphology' (p. 2), including stress assignment.

CFGs are one of a number of types of grammar formalism in the Chomsky Hierarchy (Chomsky 1959), represented in Figure 1. All of these grammar formalisms are members of the family of PSGs. The place of a particular grammar within the hierarchy is determined by the type of rules included in the grammar, as shown in Table 1 (adapted from Coleman 1998: 79).

Figure 1. The Chomsky Hierarchy

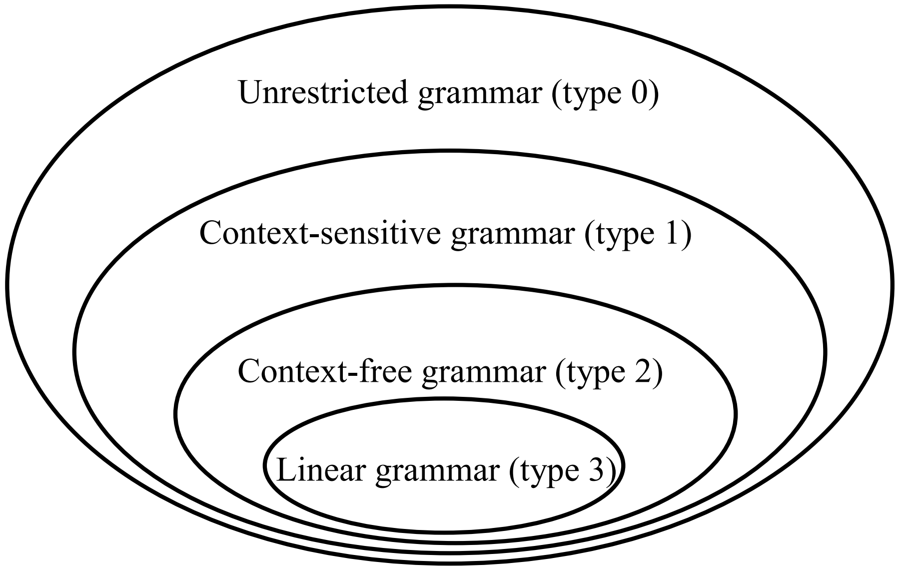

Table 1. Types of rules permitted by grammars in the Chomsky Hierarchy[3]

Type	**Grammar**	**Rule types allowed**	**Conditions on symbols**
0	Unrestricted	$A \rightarrow B$	$A \in (V_T \cup V_N)^*$ $B \in (V_T \cup V_N)^*$
1	Context-sensitive	$A \rightarrow B/ C \_ D$ and $A \rightarrow Ø$	$A \in V_N$, $B \in (V_T \cup V_N)^+$, $C,D \in (V_T \cup V_N)^*$
2	Context-free	$A \rightarrow B$	$A \in V_N$, $B \in (V_T \cup V_N)^*$
3	Right linear	$A \rightarrow aB$	$A \in V_N$, $B \in (V_N \cup \{Ø\})$, $a \in V_T$
3	Left linear	$A \rightarrow Ba$	$A \in V_N$, $B \in (V_N \cup \{Ø\})$, $a \in V_T$

[3] Note to Table 1: Following Chomsky (1959) and Coleman (1998), V_T represents the set of terminal symbols, V_N the set of non-terminal symbols, X^* a sequence of zero or more Xs, X^+ a sequence of one or more Xs, and Ø the empty string.

Because there has been a considerable amount of work carried out in phrase-structure grammar, the properties of different types of PSG in the Chomsky Hierarchy are by now well understood. These properties are important to consider when formulating a theory, for reasons which will now be made clear.

On a very general level, the more restricted the grammar formalism, the better. This follows, essentially, from the principle of Occam's razor: as Coleman (1998: 80) points out, 'the goal… in developing a formal theory of natural-language syntax or phonology, is to use a type of grammar which is as powerful as necessary, but as restrictive as possible.' It should be acknowledged, however, that context-free grammars can in practice have a cost compared to more powerful types of grammars, in that more powerful grammars may describe the same phenomena more simply (with fewer features or more general rules, for example), and may even be able to parse and generate more efficiently in some cases (Weinberg 1988).

However, there are other, perhaps more psychological, arguments in support of choosing a grammar formalism no more powerful than context-free. Bresnan and Kaplan (1982) set out a number of constraints that they suggest linguistic theory should impose on the class of possible grammars, and CFGs adhere to all but one of these constraints. The one constraint which CFG does not adhere to is the 'universality constraint', which assumes that 'the procedure for grammatical interpretation, m_G, is the same for all natural language grammars *G*' (Bresnan and Kaplan 1982: xlvii). It is significant that Bresnan and Kaplan's grounds for stating that CFG does not adhere to this constraint come from syntax, not phonology:

> Bresnan, Kaplan, Peters, and Zaenen 1982 have shown that there is no context-free phrase-structure grammar that can correctly characterize the parse trees of Dutch. The problem lies in Dutch cross-serial constructions, in which the verbs are discontinuous from the verb phrases that contain their arguments… The results of

> Bresnan, Kaplan, Peters, and Zaenen 1982 show that context-free grammars cannot provide a *universal* means of representing these phenomena. (p. xlix)

Of the other constraints, one is the *creativity constraint*. One of the claimed contributions of generative grammar to linguistics was the observation that if a grammar is to be an equally valid model both of linguistic perception and production, it should be able not only to assign structure to strings, but also to generate strings (hence the term 'generative grammar'). This observation is, for example, one of the foundational tenets of Chomsky (1957). As noted by Matthews (1974: 219), generative linguistics was partly a reaction to structuralist linguistics, which (it was claimed) emphasized assignment of structure at the expense of generation. Despite the emphasis of generative linguists upon the 'generative', it is notable that context-sensitive grammars and those more powerful are *not* necessarily reversible (Bear 1990). However, CFGs do always have the property of reversibility: that is, they can be used either for generation or recognition.

Another constraint which CFGs satisfy is Bresnan and Kaplan's *reliability constraint*: that is, they can always accept or reject strings in a finite amount of time. One of the properties of context-free (and more restricted) languages is that of decidability (alternatively known as computability, Turing-decidability or recursiveness). A language L is decidable if there is an algorithm for determining membership in L; in other words, L is decidable if there is a grammar which can decide whether a string is well- or ill-formed (a member of L or not) in a finite amount of time. Languages of type m, where $m \leq 1$, are not necessarily decidable, but those of type n, where $n > 1$, are always decidable. Bresnan and Kaplan argue that natural languages must be decidable, since:

> It is plausible to suppose that the ideal speaker can decide grammaticality by evaluating whether a candidate string is assigned (well-formed) grammatical relations or not. The syntactic mapping can thus be thought of as reliably computing whether or not any string is a well-formed sentence of a natural language. This motivates the *reliability constraint* that the syntactic mapping must provide an effectively computable characteristic function for each natural language. (p. xl)

The principal objection which has been raised to this assumption, and one which is noted by Bresnan and Kaplan, is that native speakers often do not do well at parsing 'garden path' constructions such as *The canoe floated down the river sank* and *The editor the authors the newspaper hired liked laughed*. However, they suggest, plausibly, that these constructions do not disprove their hypothesis. After all, they argue, speaker-hearers *can* disambiguate these sentences and 'recover from the garden paths', given more (but not infinite) time, and possibly a pencil and paper.

A third reason for choosing the formalism of CFG is that the ordering of the rules of CFGs will not affect the way in which they function or their end result (although the ordering of *application* of rules may have an effect on the outcome). All forms and constraints in CFGs are partial descriptions of surface representations, no rules do not ultimately constrain surface forms, all constraints must be compatible and apply equally, and any ordering of constraints will describe the same surface form (Scobbie, Coleman and Bird 1996). The motivation for this *Order-free Composition Constraint*, as Bresnan and Kaplan (1982: xlv) call it, is 'the fact that complete representations of local grammatical relations are effortlessly, fluently, and reliably constructed for arbitrary segments of sentences' (Bresnan and Kaplan 1982: xlv). Again, this does not hold for all types of grammar.

There are thus a number of reasons why it is desirable to restrict a grammar so that it is no more powerful than context-free. To summarize, these are as follows: