Structure in Language

A Dynamic Perspective

Thomas Berg

Artificial Intelligence Series

Structure in Language

Routledge Studies in Linguistics

1. Polari—The Lost Language of Gay Men
Paul Baker

2. The Linguistic Analysis of Jokes
Graeme Ritchie

3. The Irish Language in Ireland
From Goídel to Globalisation
Diarmait Mac Giolla Chríost

4. Conceptualizing Metaphors
On Charles Peirce's Marginalia
Ivan Mladenov

5. The Linguistics of Laughter
A Corpus-Assisted Study of Laughter-Talk
Alan Partington

6. The Communication of Leadership
Leadership and Metaphor Beyond the West
Jonathan Charteris-Black

7. Semantics and Pragmatics of False Friends
Pedro J. Chamizo-Domínguez

8. Style and Ideology in Translation
Latin American Writing in English
Jeremy Munday

9. Lesbian Discourses
Images of a Community
Veronika Koller

10. Structure in Language
A Dynamic Perspective
Thomas Berg

Structure in Language

A Dynamic Perspective

Thomas Berg

Routledge
Taylor & Francis Group
New York London

First published 2009
by Routledge
270 Madison Ave, New York, NY 10016

Simultaneously published in the UK
by Routledge
2 Park Square, Milton Park, Abingdon, Oxon OX14 4RN

This edition published in the Taylor & Francis e-Library, 2008.

"To purchase your own copy of this or any of Taylor & Francis or Routledge's
collection of thousands of eBooks please go to www.eBookstore.tandf.co.uk."

Routledge is an imprint of the Taylor & Francis Group, an informa business

Library of Congress Cataloging in Publication Data

Berg, Thomas, 1957–
Structure in language : a dynamic perspective / Thomas Berg.
p. cm. -- (Routledge studies in linguistics ; 10)
Includes bibliographical references and index.
ISBN 978-0-415-99135-3
1. Structural linguistics. I. Title.

P146.B39 2009
415--dc22 2008014891

ISBN 0-203-89016-7 Master e-book ISBN

ISBN10: 0-415-99135-8 (hbk)
ISBN10: 0-203-89016-7 (ebk)

ISBN13: 978-0-415-99135-3 (hbk)
ISBN13: 978-0-203-89016-5 (ebk)

Contents

For Johannes Wiermann,
the second most important non-academic teacher
to have had a profound influence
on my attitude towards life

Preface

This monograph presents a synthesis of a good portion of the research I have conducted over the past ten years or so. Although this body of work is quite diverse, ranging from historical syntax to phonological disorders, it was clear to me from the outset that a common thread ran through it all. The task I set myself in this book was therefore to string the various pieces together and to develop a unified theory that is broad enough to embrace the disparate phenomena under consideration. Viewed from a slightly different perspective, this book represents an attempt to work out the ramifications of a psycholinguistic model sketched in Berg and Abd-El-Jawad (1996) that may be seen as the embryonic form of the present monograph. The multitude of ramifications has led to the wide scope of the book with a concomitant vulnerability on all fronts. However, this was an inevitable consequence of the desire to assess the generality of the theory. This appeared all the more desirable as the fractionalization of the field makes it increasingly difficult to see the overall picture.

This brief account of the origin of the book explains (at least in part) why I had to sacrifice one of my holiest publication strategies, which is "if you end up duplicating your own work, you'd better not start publishing" (even though I am ready to acknowledge that monographs follow a somewhat different logic from journal articles). Because it was my overall aim to bring together the various strands of research under the umbrella of a single theory, I could not help quoting my previous publications. In particular, section 2.4.1 is a modified version of Berg (2003b), section 4.2.1 follows Berg (2002b) closely, section 8.2.2 summarizes Berg (2006) and section 9.2 relies on Berg (2002a) as well as Berg (1997). The relevant parts are reproduced by kind permission of Cambridge University Press, Elsevier, and Peter Lang. Some of the ideas contained in this book were first submitted to audiences in Freiburg, Paderborn, Berlin (all in 2001), Boston, MA in 2005, and Bremen in 2006.

Numerous people have contributed in one way of another to this monograph. I was fortunate to receive comments on the text from Winfried Boeder, Florian Dolberg, Ulrich Schade, and Peter Siemund and above all from the anonymous Routledge readers. Nigel Isle has been my faithful

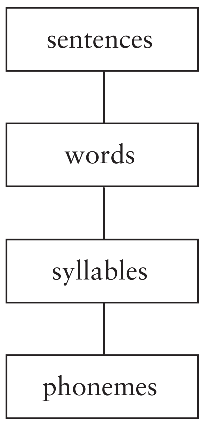
companion on the long road to linguistic correctness. Kath Baker shared her native-speaker intuition with me more than once. The book's long gestation period has seen quite a few student assistants and some others who were of invaluable help in the data collection process, to wit Beata Zaide, Andreas Sohr, Sabine Helmer, Christian Koops, and Anatol Stefanowitsch, or who were immensely serviceable in turning the manuscript into its final shape, viz. Ole Christiansen, Sandra Lund, Marion Neubauer, and Maren Schiefke. The experiment reported in Chapter 10 (and its mute forerunner) would not have been possible without the help of Rik Eshuis, Magdalene Emmelius, Trevor Harley, and Beth Wilson. I could always count on Stefan Gries's expertise in statistical matters. My heartfelt thanks go to all of them.

1 A Structural Model of Language Production

1.1 THE DOUBLE HIERARCHY

One of the incontrovertible facts of language is its hierarchical organization. Although the number of levels and their relationship to one another may be a matter of dispute, there is general consensus that linguistic units are amenable to a ranking by size, as illustrated in (1).

(1) The hierarchical organization of some major linguistic units



It is equally uncontroversial, although not widely acknowledged, that the units in (1) differ in their psycholinguistic status. Whereas some units are "there," like books on a shelf, waiting to be taken and used, others are not "there" and therefore have to be constructed.

Monomorphemic words and phonemes are clear instances of the former category. A speaker must have a repository of lexical items and these items must provide permanent access to their phonological representation, which includes information about the nature of the constituent phonemes and their order. A successful use of language thus requires that monomorphemic words and phonemes be part of the long-term memory representation of the ordinary speaker.

The case of sentences is different. It is obviously true that the infinite number of different sentences, whether actually attested or potentially constructible in compliance with the rules of the language, stands in the way of committing them to memory.[1] We must therefore take it as an established fact that sentences cannot normally be retrieved as holistic units but have to be built up during the preparation to speak.[2]

The case of syllables is different again. We know that the number of syllables in a language is finite and considerably lower than that of words (Schiller, Meyer, Baayen, & Levelt, 1996). The arguments of creativity and limited storage space that apply to syntax thus cannot be extended to phonology. However, this does not mean that syllables have the same status as words and phonemes. The major linguistic criterion that is employed in this connection is that of redundancy. If a particular unit is redundant (i.e., predictable from some other source of information) it need not be stored in the lexicon. Given that syllable boundaries are not normally distinctive,[3] there is no motivation for including syllables in the long-term memory representation. Consequently, syllables would have to be actively put together much like sentences, though for very different reasons.

We thus arrive at an initial division of linguistic units into two classes—those that can be taken off the shelf ready-made ("prefabs") and those that have to be created in an ad hoc manner ("assemblemes"), as shown in (2).

(2) Two sets of linguistic units

Ready-made units	Ad hoc units
monomorphemic words	sentences
phonemes	syllables

These two sets of linguistic elements make up two different hierarchies that must be systematically related to each other to ensure the smooth production of language. (This aspect is not expressed in (2).) The ready-made elements will be termed *content units* and the ad hoc elements *structural units*. It is the purpose of this book to work out the implications of this distinction for a general theory of language—general in the sense that it aims at covering major aspects of language structure, change, acquisition, and loss. The focus of this enterprise will be on the structural side of the coin, which will be systematically explored. We begin by motivating the terms *content* and *structure* and setting them off against other prevalent uses in the relevant literature. This is followed by an assignment of linguistic units to either the content or the structural domain.

1.2 CONTENT VERSUS STRUCTURAL UNITS

Content units have just been defined as being available in a speaker's long-term store, unlike structural units that have to be made up on the spot. This definition of structure is certainly not standard. In fact, it represents quite a radical departure from previous notions of structure. At least three different concepts are discernible in the linguistic literature. The first originates with the structuralists of the first half of the last century (e.g., Hill, 1958) and is still current today. This notion by and large equates structure with hierarchy. In this view, every linguistic element is of the structural kind so long as it is part of a larger chunk. As language is hierarchically organized, the implication is that all linguistic elements are structural in nature. The following statement from Coates (1999, p. 2) is entirely typical: ". . . lots of words . . . do evidently consist of smaller pieces—they have STRUCTURE" (emphasis his). By "smaller pieces," Coates means morphemes and by "structure," he means internal structure. He therefore regards morphemes as structural units. In a very similar vein, Greenberg (1957, p. 81) sees phonemes as "substructures." It goes without saying that this view of structure is incompatible with the distinction between content and structural units.

On the second reading, structure is short-hand for syntactic structure. This equation of structure with syntactic structure probably stems from the fact that syntax is widely recognized as the structural domain par excellence. Whereas the first definition of structure is too general in the light of the content–structure distinction, this one is too narrow in that it assigns syntax a uniqueness it may not have.

The third perspective on structure defines it in opposition to meaning. A pertinent example is the concept of boundary, in particular the distinction between word boundaries (#) and morpheme boundaries (+) introduced by Chomsky and Halle (1968). As Aronoff (1976, p. 121) puts it, ". . . boundaries are structural entities. . . . Like all structural entities, they have no phonological substance in themselves, nor meanings in the conventional sense . . .". It is obvious from this quotation that structure is something invisible or inaudible, a type of unit that lacks both signifier and signified. This view, then, stands in maximum contrast to the first conception of structure. Whereas all of the traditional linguistic units are regarded as structural on the first reading, none of them possesses this status on the last reading because they all have either a signifier or both a signifier and a signified. This third definition of structure contrasts sharply with the idea of sentences and syllables as structural units and leaves no room for the content–structure distinction more generally.

In view of the wide array of meanings attributed to the term *structure*, it may seem unwise to use it in yet another sense. However, what motivates the use of this term in the present context is its etymology, which conveys precisely the idea that puts us on the right track (even though, as will be seen below, it requires considerable elaboration). The term *structure* derives from the Latin noun *structura*, which in turn comes from the verb *struere*, whose original meaning was largely confined to the construction of buildings. Such work involves the assembly of smaller parts to create larger ones, and in fact, this notion is clearly embodied in the term *structure* in both its ordinary use (e.g., a six-story concrete structure) and its scientific use (compare e.g., Hartmann, 1964; Wunderlich, 1971).

The term *content* poses less of a problem because it conflicts less with other uses around. A major way in which this term is understood is in contrast to the notion of frame. Azuma (1993), for example, works with a model that provides for a syntactic frame into which linguistic elements (i.e., words) are inserted. Janssen, Roelofs, and Levelt (2002, 2004) argue for morphological frames and MacNeilage (1998) for phonological frames. These slots are filled with linguistic units such as morphemes and words. We will define here that all elements that are inserted into a frame are content units.

Another important sense in which the term *content* is used in the linguistic literature is in opposition to function. The distinction between content and function elements holds at the lexical level and is built on the semanticity–syntacticity contrast (among other criteria). Content words (e.g., nouns) are characterized by a high degree of semanticity, whereas function words (e.g., the infinitival particle *to*) are syntactically motivated. In this sense, then, content is synonymous with lexical meaning.

The way in which the term *content* is used here is much more congenial to the former than to the latter usage. The latter definition is too narrow and does not capture some fundamental similarities between units with a signified and those without. By contrast, the former definition has essentially (albeit not completely) the same extension as the one proposed below, even though the underlying motivations are quite disparate. The state of having an entry in the mental lexicon is definitely not the same as the quality of being inserted into a frame. It is mainly this considerable overlap in the extensions of the two definitions that justifies the use of the term *content* in the ensuing analysis.

1.3 COHESIVENESS AS THE LITMUS TEST

The first challenge is to divide linguistic elements into either content or structural units. The critical question is this: Which elements are stored in long-term memory and which are not? It should be made quite clear at the outset that both types of unit are eventually "there," only their psycholinguistic history is different. Let us take as an example the units *word* and *sentence*, the former of which was provisionally assigned to the content set and the latter to the structural set in section 1.1. Clearly, both words and sentences are produced in the act of speaking. So, as products, they are both "there." However, while the content units exist from the beginning, the structural units have to be brought into existence. Of central importance is the claim that this generating takes time (i.e., it is a real-time process that begins with the absence of structure and ends with the presence of structural units; see Berg & Abd-El-Jawad, 1996, and Bertinetto, 2001a).[4] In between these limits, structural representations are gradually erected. Theoretically, this erection process continues until a structural unit has been built up completely. What does it mean for a unit to be completed? If a structural element has been created to the fullest, it acquires the status of a unit in the sense that it perfectly unites the elements it is composed of. Differently put, it has reached a maximum degree of cohesion. Cohesiveness thus becomes the foremost index of the gradual build-up of structural information. In the early phases of this process, little cohesiveness is expected, whereas later stages predictably generate more cohesiveness.

Unless the unlikely assumption is made that the head start of the content units is lost to the structural units, the former will display a greater degree of cohesion than the latter at any moment in the process of preparing to speak. Degree of cohesiveness thus turns out to be the major criterion for distinguishing between content and structural units. This leads us to the following identification procedure. An element that behaves cohesively is identified as a content unit while an element that exhibits a less than cohesive behaviour is identified as a structural unit. Note that the degree of cohesiveness is variable both across different elements and within one and the same element. Due to their inherent differences in size, make-up, and function, different structural units may differ in the time they take to reach a certain level of cohesiveness. Assuming a threshold at which the erection process is stopped and articulation begins, different structural units may attain different degrees of cohesion. Moreover, one and the same unit need not always be equally cohesive. If it can be “caught” at different moments in the erection process, it will evince different degrees of cohesion. By averaging across all relevant data points, we may derive a general measure of the cohesiveness of a particular unit. Since we are talking about degrees of cohesiveness, a sharp dividing line can be expected neither between content and structural units nor between structural units and no units at all. It is theoretically possible for a structural unit to be fairly cohesive and thereby approximate to the behaviour of content units. Similarly, if a structural element is highly incohesive, it behaves almost as if it was not there and accordingly may be difficult to make out.

In the following discussion, a wide range of linguistic units will be examined in terms of their membership of the content or structural group. Although the distinction between content and structural units is assumed to be universal, the analysis will be performed using data mainly from English. The list of units is not meant to be exhaustive, although the lower levels are covered more extensively than the higher ones. To determine the cohesiveness of these units, empirical data are necessary that highlight certain parts of an utterance against the background of the utterance as a whole and that may arise at different moments in time in the language-planning process. These two requirements are perfectly met by speech errors (i.e., inadvertent deviations from the speaker’s intention). A further great advantage of this data type is that it is uncontaminated by speakers’ preconceptions and experimenters' instructions among other potential distortions. Speech errors are local phenomena, that is, they pick out individual units from their surrounding context and leave the remainder untouched. In addition, the least contentious hypothesis holds that slips of the tongue may occur anytime in the language generation process because there is no reason to assume that some stage in the production process is immune to malfunction while another is not. In point of fact, empirical evidence in support of the claim that errors may arise at various temporally defined points in the generation process has been accumulating over the years (e.g., Stemberger & Lewis, 1986, Berg, 1992a; Berg & Abd-El-Jawad, 1996).

The order in which the data will be presented is from larger to smaller units. One of the notable discoveries of speech error research is that the largest units are rather small. Indeed, the largest units involve no more than two words simultaneously. Two such cases are documented next in (3) and (4). All utterances appear in the sequence in which they were produced (i.e., the erroneous utterance precedes the target utterance, either actually produced by the perpetrator or reconstructed by the error collector). The error units are italicized for ease of identification.

(3) If you'll *meet him* you'll *stick around*. for: If you'll stick around you'll meet him. (from Fromkin, 1973)

(4) I'd like to speak to *this matter* about *you*. for: I'd like to speak to you about this matter. (from Fromkin, 1973)

Case (3) involves a reversal of two VP's, and (4) a reversal of two NP's. The fact that one of the interacting units in (3) is a phrasal verb is certainly not coincidental, as phrasal verbs are characterized by a high degree of idiomatization. This aspect implies a high degree of cohesiveness at the semantic level, which is unlikely to be completely lost during the translation process from a semantic to a syntactic representation. It is also not surprising to observe some cohesion between *this* and *matter* in (4) because the determiner modifies the noun. It might be fitting to add the obvious fact that the two-word units misordered in (3) and (4) are adjacent. The first conclusion to be drawn is that the units involved in two-word errors are semantically and syntactically well-defined.

However, much more important in the present connection is the uncommonness of these two-word errors. Whatever the syntactic relationship a word may have contracted, it most usually is affected individually, as in (5) which in a sense is the counterpart to (4).

(5) a small body of *instruments* written for these *compositions*.
for: a small body of *compositions* written for these *instruments*.
(from Fromkin, 1973)

In (5), the NP *these compositions* is broken up. Since this is the typical case, it may be concluded that complex NPs generally display a low degree of cohesiveness. Generalizing, we may go so far as to claim that all syntactic phrases consisting of at least two words rank fairly low on the cohesiveness scale. This result allows us to confirm what was claimed earlier for sentences. On the cohesiveness criterion, both sentences and syntactic constituents (above the single-word level) are structural units.

For expository reasons, the next unit to be investigated is the monomorphemic word, even though polymorphemic words are obviously larger in size. By word we understand a free-standing unit in contrast to a morpheme, which is defined here as bound in order to avoid the classificational ambiguity of items such as *brick*, which are generally viewed both as a word and as a morpheme. Monomorphemic words display a very clear pattern. They are one of the most frequent error units (i.e., they act holistically in the error process). A standard example is provided here.

(6) Although *murder* is a form of *suicide*. for: Although suicide is a form of murder. (from Garrett, 1975)

Case (6) exemplifies the reversal of two words that are misplaced as wholes. Because this example is entirely typical, the cohesiveness of monomorphemic words is beyond doubt. They are accordingly assigned to the category of content units.

An entirely parallel behaviour can be observed with morphemes, no matter whether they are lexical or grammatical, stems or affixes, prefixes or suffixes. Whenever they are involved in a malfunction, they act as units, as in (7).

(7) You want the potatoes *slice*ly *thin*ned? for: thinly sliced? (from Stemberger, 1985)

This is a typical morpheme error in which the lexical morphemes *thin* and *slice* exchange places. Although the morphemes are part of larger units, they themselves do not disintegrate in the error process, thereby testifying to their cohesiveness. By implication, they will be treated as content units.

Error (7) leads us to a consideration of polymorphemic words. As this slip shows, the words *thinly* and *sliced* are split, and their morphemes are individually affected. Thus, they are less than cohesive. How representative is this example? When examining the cohesiveness of morphologically complex words, it is useful to follow the standard practice of distinguishing between inflected and derived words as well as compounds.

Let us begin with inflected words. The following two examples illustrate the basic decision that the processor has to make in dealing with morphologically complex words.

(8) Well you can cut *rain* in the *trees*. for: trees in the rain. (from Garrett, 1982)

(9) He doesn't have any *closet*s in his *skeleton*. for: skeletons in his closet. (from Stemberger, 1985)

The two slips display a massive parallelism of sentence structure. They exemplify a sequencing problem between a singular and a plural noun. There are two options. Either the plural marker accompanies its misordered host or it stays behind and attaches to its new host. The former alternative is documented in (8), the latter in (9). It is quite evident from Stemberger's (1985) database that (8) is the exception and (9) the rule. He has 135 pertinent errors in his corpus of which 120 (= 88.9%) leave the inflection behind and 15 (= 11.1%) take it with them. This low degree of cohesiveness leaves no doubt that inflected words are not normally stored in the mental lexicon (Stemberger & MacWhinney, 1988). They are structural units. It might be added that the percentage of whole-word errors varies with the type of inflection and that it is even lower for most other inflections than for the plural. We thus have to reckon with slightly different degrees of cohesiveness for different inflections.

The next step leads us to consider compounds. As these are less common in English than in German and as quantitative information on English compound errors is not available, I will dip into my own collection of German slips of the tongue. Compounds may be implicated in errors in three different ways. A compound may interact with another compound, a compound may substitute for a noncompound, or vice versa. The first case is illustrated by the following examples, which are augmented by interlinear glosses and translations.

(10) Gestern hat die chemische Industrie auf der *Pressekonferenz*—auf der *Hannovermesse* eine Pressekonferenz gegeben.
yesterday has the chemical industry at the press conference—at the Hanover Fair a press conference given
'Yesterday the chemical industry held a press conference at the Hanover Fair.'

(11) Wir haben morgen Elternabend vom Kinder*abend*—vom Kinder*garten*.
we have tomorrow parents evening of the children evening—of the kindergarten
'Tomorrow we will have a parents' meeting of the kindergarten.'

Again, the parallelism of the two slips allows us to study the basic problem that the processor faces. Either the (bilexemic) compound is replaced in full or only one lexeme is replaced while the other stays put. The former happened in (10), the latter in (11). The compound *Pressekonferenz* (press conference) is anticipated to replace *Hannovermesse* (Hanover Fair) in (10) while the lexeme *Abend* (evening) from the compound *Elternabend* (parents' evening) perseverates to replace the lexeme *Garten* (garden) from the compound *Kindergarten* (kindergarten).

After the exclusion of all ambiguous cases, a total of 46 relevant errors were found in the German database. Of these, 37 (= 80.4%) underwent splitting (as in (11)) and 9 (= 19.6%) were of the holistic type (as in (10)). This asymmetry holds equally for all three subsets of compound errors just mentioned. These figures invite the conclusion that compounds exhibit a rather low degree of cohesiveness (see also Blanken, 2000, for German, Badecker, & Caramazza, 1998, and Badecker, 2001, for English compounds). They are therefore best viewed as structural units. In other words, this approach stresses the compositional nature of compounds.

The same procedure was applied to derived words. The opposite ways of treating this set of morphologically complex words are shown below.

(12) Is there a cigarette *building* in this *machine*? for: a cigarette machine in this building. (from Garrett, 1980)

(13) Can I have a full *nud*al *front*ity? for: a full frontal nudity. (from Stemberger, 1985)

The bimorphemic word *building* interacts in toto with the monomorphemic word *machine* in (12). However, the derivational suffixes *–al* and *–ity* remain in their original location and attach to their new lexemes in (13). In stark contrast to what was found for inflected words and compounds, Stemberger (1985) observes that the majority of derived words in English act in unison. There are 12 relevant slips in his sample of which 9 (= 75%) leave the complex word intact and 3 (= 25%) break it up. Despite the low number of mistakes, it may safely be concluded that derived words rank much higher on the cohesiveness scale than inflected words and compounds.

Translating the 75% cohesiveness index into the binary opposition of content and structural units turns out to be a difficult undertaking. As was pointed out towards the beginning of this section, the general approach adopted here does not lead one to expect a clear-cut distinction between the two types of units. Although one might easily be fooled into believing in the dichotomous nature of linguistic units—they are either stored in, or missing from, long-term memory—it appears much more appropriate to envision a complementary relationship between content and structural units. We could either say that a derived word is basically a content unit that is backed up by structural information or that it is a structural unit that is strongly backed up by content information. Although the first option is bolstered by the fact that the majority of derived words exhibit a cohesive behaviour, preference will, however, be given to the second option. Because content units were defined as being generally cohesive, any element that fails to reach this criterion cannot be assigned to the same group. A 25% break-up rate thus seems reason enough to identify derived words as structural (i.e., compositional rather than holistic) units. This classification takes account of the differential behaviour of monomorphemic and polymorphemic words. Furthermore, it assigns all polymorphemic words, whether inflectional or derivational, to the same category and thereby emphasizes the fundamental similarity between them, without, however, negating their differences. These may be attributed to the varying impact of long-term memory on the production of morphologically complex words. Other factors such as function and frequency also have to be considered in this context.

We now proceed to the levels below the word. The largest element below the word level is what Berg (1989a) termed the *superrime*. It consists of a rime plus a full (unstressed) syllable. Consider (14).

(14) str*unction* and f*ucture*. for: structure and function. (from Garrett, 1975)

Number (14) involves an exchange of two disyllabic words minus their initial consonants. It thus evidences a break point between the initial consonant(s) and the remainder of the word, viz. the superrime. It is a fact that superrime slips are uncommon. The by far more common error type is the whole-word slip as the closest alternative. The low cohesiveness of superrimes makes it quite clear that they belong to the category of structural units.

The constituents of the superrime lead us directly to the analysis of the syllable. We owe to Shattuck (1975) the baffling discovery that syllables are only very rarely implicated in slips of the tongue. One of the few uncontroversial cases is in (15).

(15) guitune my tar. for: tune my guitar. (from Shattuck-Hufnagel, 1979)

This slip exemplifies a leftward shift of the syllable /gi:/, which docks onto the word *tune* to produce *guitune*. The rarity of (syntagmatic) syllable errors shows that the cohesiveness of syllables is very low. Therefore, they are unhesitatingly classed as structural units. It is worth noting that the cohesiveness argument gives exactly the same result as the redundancy argument resorted to in section 1.1.

We now leave the syllable for its constituents, in particular the rime that encompasses everything from the vowel to the coda consonant(s). Refer to (16).

(16) The juice is still on the table. Is that en*uice*? for: en*ough*. (from Garrett, 1980)

This error shows a perseveration of the rime in *juice*, which intrudes on the rime of the stressed syllable in *enough* to yield *enuice*. Rime errors are relatively infrequent, and certainly less frequent than single-phoneme slips, which testifies to the low degree of cohesiveness of rimes. Like syllables, they unequivocally qualify as structural units.

Rimes represent a combination of minimally a vowel and a consonant. The combination of two consonants, that is consonant clusters, will be examined next. There are two types of clusters—tautosyllabic and heterosyllabic. We focus on the former type as it befits the notion of a hierarchy to study elements within elements, not across elements. Clusters come in different subsets depending on the nature (and number) of adjacent consonants. A typical representative of the cluster category is the obstruent + liquid set on the basis of which the contrast between a holistic and an analytic treatment will be illustrated.

(17) *c*oat *thr*utting. for: throat cutting. (from Fromkin, 1973)

(18) *th*eep *d*roat. for: deep throat. (from Shattuck-Hufnagel, 1983)

Examples (17) and (18) present a rare pair of errors in which one and the same word undergoes a differential treatment. The cluster /θr/ in the word *throat* acts as a unit in (17) but splits in (18). Note that there are no phonotactic reasons for this divergent behaviour. The cluster /θr/ interacts with /k/ in (17) and /d/ in (18). Both these consonants can readily be followed by a rhotic, the phoneme sequence that would have resulted if the initial consonant alone had been dislocated (as in (18)).

Berg (1994a) provides a quantitative analysis of the cohesiveness of initial stop + sonorant clusters in English. Of 109 pertinent slips, 85 (= 78.0%) break the cluster up (as in (18)) while 24 (= 22.0%) leave it unscathed. There is hardly any difference in cohesiveness between the various cluster subsets such as /Cl-/ and /Cr-/, apart from the fact that /tr/ and /dr/ prove to be more cohesive than the other stop + rhotic clusters. However, the inverted-sonority type /s/ + stop behaves quite differently. Berg (1994a) reports that clusters like /st/ for example, stick together in 69% of the relevant slips. It may be inferred from these results that the cohesiveness of a cluster depends on the phonological class it belongs to. While /s/ + stop clusters are fairly cohesive, obstruent + sonorant clusters usually fall apart in the error process. Despite these differences in cohesiveness, it may be submitted that clusters in general are structural units. This hypothesis is rather uncontroversial in the case of obstruents + sonorant clusters. The assumption that it also holds for /s/ + stop clusters allows us to account for the (relatively few) incohesive cases. The cohesiveness of these clusters requires a different explanation, presumably one in terms of sonority relationships among their constituents. It might be worth mentioning as an afterthought that heterosyllabic clusters are even less cohesive than tautosyllabic ones.

We have now reached the level of the single segment. It has just been shown that phonemes as parts of clusters tend to migrate individually rather than in conjunction with other consonants. Do phonemes also outstrip their smaller competitors (i.e., features)? The two levels at which speech errors may occur are exemplified in (19) and (20).

(19) Syllable re*s*er*v*als do occur. for: syllable reversals. (from Trevor Harley, unpublished)

(20) *z*eek *f*erification. for: seek verification. (from Fromkin, 1973)

Both (19) and (20) document a problem of ordering the two phonemes /s/ and /v/. In (19), their integrity is preserved during misordering. In (20), by contrast, the intended fricatives /s/ and /v/ turn up as /z/ and /f/, respectively. That is, this reversal took place at the feature level. In particular, the voice feature was exchanged such that /s/ adopted the [voiced] value from /v/ and /v/ the [voiceless] value from /s/. In the latter case, then, the integrity of the interacting phonemes was destroyed.

There is a good deal of agreement in the speech error literature that phonemes are cohesive units, even though it is true that many phoneme slips look ambiguous on the surface. Take (19) as an example. Theoretically, it may be construed not only as a whole-segment slip but also as a feature slip in which the place-of-articulation values of /s/ and /v/ (i.e. [alveolar] and [labial]) traded places whereas the other feature dimensions (i.e., voice and manner of articulation) were left untouched. This interpretation can be shown to be fallacious on a number of empirical and theoretical grounds, the details of which need not concern us here. The main point is that unambiguous feature errors are truly exceptional (e.g., Shattuck-Hufnagel, 1983). On this argument alone, it is justified to contend that phonemes are highly cohesive in character. Implicationally, they form a subset of the content units.

We are thus left with the phonological features themselves. As these constitute the smallest elements in the linguistic hierarchy, they by definition cannot disintegrate. They must therefore be content units.

This completes our *tour d'horizon* of the linguistic units that may be involved in English (and German) slips of the tongue. It will not have escaped the reader's notice that two phonological units have not been mentioned—the foot and the mora. The reason is simply that foot and mora errors do not seem to occur. This is certainly not unexpected in the case of feet. In the light of the fact that syllable errors are so uncommon, it comes as no surprise that sequences of two or more syllables resist misordering. All that can be said at this point is that the foregoing analysis did not yield any evidence in favour of feet.

There are a handful of slips of the tongue that might be mistaken for mora errors. Look at (21).

(21) *ca*ssy *pu*t. for: pussy cat. (from Fromkin, 1971)

At the descriptive level, (21) involves the permutation of the phoneme sequences /pʊ/ and /kæ/. On one version of moraic theory, which is espoused by, for example, Hyman (1985), the syllable-initial consonant and the following vowel are dominated by the same mora node whereas the syllable-final consonant is dominated by a separate mora node. Hence, the link between the prevocalic consonant and the vowel is stronger than that between the vowel and the postvocalic consonant. Assuming that moras form a level of representation that is called on during the language production process, the prediction would be that CV errors occur more frequently than VC errors. However, the opposite is true (see Stemberger, 1983a and Chapter Two, this volume). This version of moraic theory should therefore be rejected and the error in (21) not be categorized as a moraic one.

The other version of moraic theory (Hayes, 1989) adjoins the prevocalic consonant to the syllable rather than the mora node. The vowel and the postvocalic consonant are moraic as in the other version. A characteristic trait of this theory is that moras do not branch. This implies that the structure of the syllable is essentially flat and therefore differences in cohesiveness between adjacent phonemes are not predicted. However, as just noted, CV and VC sequences are unequal in their cohesiveness. Consequently, this version of moraic theory also fails. The conclusion to be drawn from this is that the speech error evidence argues against the reality of a moraic level of representation.

Also missing from this discussion are other units of the prosodic hierarchy, in particular phonological words and phrases (Nespor & Vogel, 1986). The reason for this absence is not that these groupings do not exist but rather that the evidence for them is more indirect than the speech-error argument that has been made in this discussion. As the units involved in malfunctions hardly ever go beyond the single-word level (see earlier discussion), there is no way of arguing the case of larger prosodic groupings within the framework of the preceding analysis. It is clear, however, that if phonological words and phrases are real, they must be structural units. The same goes for moras and feet. If other lines of evidence should find them necessary components of the language production process, their absence in speech errors would certainly militate against their assignment to the category of content units.

A final omission is the issue of segment structure. The basic assumption is that the feature dimensions that are constitutive of a phoneme (e.g., place and manner of articulation and voice in the case of consonants) do not form a linear, unordered set but are hierarchically organized into various levels (e.g., Clements, 1985; Odden, 1991). In a nutshell, phonemes are claimed to have constituent structure not unlike that of sentences. Although it is not my intention to take issue with the theoretical phonologists' proposals as *linguistic* constructs, there is a psycholinguistic sense in which sentence structure and segment structure are fundamentally different. As will be argued in section 1.5, structure is intimately tied to the serialization of language. It helps in the fluent production of a sequence of smaller units within a larger unit. There is a major disparity between sentence and segment structure in terms of serialization. Whereas the words within a sentence have to be put into a certain order, the features within a segment must be simultaneously available. For the production of a segment, all its features must be accessed in parallel. A serial relationship among them would be fatal. It follows from this that structure in the sense used here is not only unnecessary but even detrimental at the subsegmental level. It is consequently maintained that segment structure as a psycholinguistically relevant notion does not exist.[5]

It is time to take stock of what we have discussed up to now. The concept of cohesiveness has been utilized as a diagnostic of whether a linguistic element belongs to the class of content units or to that of structural units. The results of this investigation are summarized in (22).

(22) The inventory of content and structural units

Content units	Structural units
monomorphemic words	sentences
morphemes	syntactic phrases
phonemes	compounds
features	derived words
	inflected words
	superrimes
	syllables
	rimes
	clusters

1.4 TWO TYPES OF STRUCTURAL UNITS

The study of speech errors reveals another class of units that are needed in an adequate description of the language production process. Consider the exchange in (23).

(23) Hel*f*, hel*f*, the wol*p* is after me! for: Help, help, the wolf is after me! (from Garrett, 1980)

It was Shattuck-Hufnagel (1979, p. 303ff) who was the first to ask exactly the right question: How does the /p/ that was driven out of position by the /f/ find the position that was originally inhabited by the /f/? If the /f/ had left no trace of its original position, the /p/ would have been hard put to end up in the position it actually does. Shattuck-Hufnagel's solution was to argue that the /p/, on being dislocated, left a vacant slot behind that could then be filled by the /f/. She thus postulated a two-level representation consisting of slots and fillers that have to be associated with each other in the production process. This separation of representational levels roughly corresponds to the distinction between the skeleton and the melody tier in theoretical phonology (Clements & Keyser, 1983). The melodic tier makes available a set of segments, whereas the skeleton tier generates the place-holders for consonants (C) and vowels (V) that accommodate the segments. The association between fillers and slots is usually one-to-one, but it can also be many-to-one and one-to-many. An example of all three types of association is given in (24) for the word *peach*.

(24)	skeleton tier	C	V	V	C	
	melody tier	p	i:		t	ʃ

As a general rule, all consonants including affricates are linked to a single C-slot, short vowels to a single V-slot and long vowels and diphthongs to a double V-slot. (In languages with distinctive consonant length, short consonants are adjoined to a single C-slot and long consonants to a double C-slot.) These departures from the principle of one-to-one association add a new dimension to the CV tier. Although it functions merely as a set of lined-up positions in a model that allows only one-to-one associations, the introduction of one-to-many associations makes the CV tier code quantitative information about segments. This amounts to a representational split whereby the qualitative properties of a segment are coded at the melodic and the quantitative properties at the skeleton tier.

This representational segregation generates an interesting prediction about the behaviour of quantitative and qualitative aspects of phonemes. As the two are represented at different levels and as each level may be reasonably assumed to be affected individually in the error process, one would expect quantitative information to be separated from qualitative information in slips of the tongue. Precisely this happened in (25), an example from German, a language with a consistent vowel length contrast.

(25) M*i*ll- Melanie.
[mil mɛlani:]

This error shows an interaction between the long vowel /i:/ and the short vowel /ɛ/. As can be seen, the word-final /i:/ is "oblivious" to its length specification and surfaces as /i/, that is, it adopts the length of the vowel it replaces (/ɛ/). This metamorphosis is readily explained on the assumption that this slip occurred at the melodic tier and ignored the skeleton tier. The fact that such dissociations between quantity and quality are the rule provides strong support for the dual-representation hypothesis in (24).

What is the status of the CV tier? To be more specific, is it stored in long-term memory? Roelofs & Meyer (1998) deny that the CV structure is included in the permanent representation of lexical items because this level failed to make an impact in their priming experiments. However, Meijer (1996) obtained facilitatory effects due to similarity at the skeleton tier (see also Ferrand & Segui, 1998; Costa & Sebastian-Gallés, 1998; and Berent, Bouissa, & Tuller, 2001) and consequently took his data to mean that the CV structure is stored with each lexical item. In a similar vein, Prunet, Béland, and Idrissi (2000) argued that the representation at the skeleton tier is memorized because the linking between the skeleton and the melody codes serial-order information that cannot be computed by rule.

There are problems with the latter view. It does not seem defensible to delegate serial-order information to the CV tier or to the associative links between the skeleton and the melody tier. This information must be included in the melody tier as the speaker's task is not to put a series of C and V units into proper order but a series of individual phonemes. The links also cannot be held responsible for serial ordering because, without prior knowledge about the correct order, the association process cannot operate smoothly.

The real problem with the CV tier is that it is implicitly assumed to fulfil several functions simultaneously. In order to determine the psycholinguistic status of the CV tier, it is necessary to keep these functions apart. It is certain that the skeleton's function of coding quantitative information is part of a word's permanent representation. The length of a given segment or the number of times it must be produced is an idiosyncratic property of lexical items. It is as unpredictable as the quality of a phoneme. The best proof for this claim is the distinctive nature of quantitative information. For example, vowel length is the distinguishing phonological trait in the German minimal pair *Miete* [miːtə] 'rent' versus *Mitte* [mitə] 'middle.'

By contrast, the skeleton's function of providing for place-holders need not be included in long-term memory. If each segment is assigned one slot, the CV structure is entirely redundant. That is, there would be no need to clutter up one's memory with it.[6] The same is true of the skeleton's final function of coding major-class information (i.e., the distinction between consonants and vowels). Because this information is predictable from the phonemes themselves, there is no reason to have a CV tier on which it is permanently represented.

The conclusion seems inevitable that the provision for slots and the coding of quantitative information are functions of the skeleton tier that should be kept representationally distinct. It is necessary therefore to create an additional representational level and assign these two functions to distinct levels. Specifically, the function of providing slots will be reserved to what will henceforth be called the *slot level* whereas the function of coding quantitative information will be reserved to what we will dub the *quantity level*. For reasons of terminological consistency, the melody tier will be renamed the *quality level*. Of course, it makes no sense to represent the consonant–vowel distinction at the quantity level. This information is therefore located at the slot level.

The units at the three levels change their names according to their new definitions. The quantity level knows only the numbers 1 and 2 (i.e., indications of how often a unit at the quality tier is to be produced). This way, doubling information is no longer represented by associating a geminate unit with two positions at the CV tier, as in other models, but rather by associating it with a doubling marker that has no other function but to code quantity. This suggestion is quite similar in spirit to an idea briefly expressed in Miceli, Benvegnú, Capasso, & Caramazza (1995). The slot level is endowed with icons symbolizing containers (∪). No changes are necessary at the qualitative level.

As regards the organization of the three levels, the quality level takes the place between the quantity and the slot level. It is obvious that the quantity and the quality level must be adjacent. It is also clear that the slot level links up with the quality rather than the quantity level because the former, though not the latter, provides the fillers that go into the slots. Note that there is a consistent one-to-one correspondence both between the elements at the quality and the slot level and those at the quality and the quantity level.

Diagram (24) can now be expanded into (26).

(26) A three-tiered representation of the word *peach*

quantity level	1	2	1
	|	|	|
quality level	p	i	tʃ
	|	|	|
slot level	∪	∪	∪

This three-tiered model preserves the strengths of the former two-tiered model but has the additional advantages of separating functions that are logically independent and of distinguishing levels that have a different psycholinguistic status.[7] As explained earlier, the quantity tier must be part of an item's long-term memory representation. However, the slot level is denied a place in long-term memory. By the definition set out in the first section, slots belong to the set of structural units.

It is worth noting that the three-tiered representation in (26) is mainly confined to the phonological component because quantity is largely a non-issue in other domains. In syntax, the problem does not arise as there are no lexical representations for sentences or parts thereof (see section 1.1). An ADJP such as *very, very useful* is therefore generated by accessing the lexical node for the intensifier twice. Morphological complexes may have a (weak) lexical representation (see section 1.3), but they are not normally distinguished by a difference in the number of identical morphemes. For example, there is no lexical opposition between *childhood* and **childhood-hood*. The only exception appears to be morpheme-based reduplication. Two types have to be distinguished. The reduplicated form may consist of units that do not exist on their own such as *wakey-wakey*. Because the morphemic status of these units is uncertain, it is not clear whether the reduplicative process belongs in the morphology rather than the phonology. In any event, these cases are rather uncommon. More frequent are reduplications in which a true morpheme or word is needed twice, such as *hush-hush* and *buddy-buddy*. Here, the quantitative information is distinctive much as vowel length in the aforementioned pair *Miete* and *Mitte*. Therefore, the following representation appears justified.

(27)	quantity level	1	'hush'	2	'hush-hush'
		|		|	
	quality level	hush		hush	

It is entirely reasonable to assume that the quantity level remains unspecified in all cases where a given morpheme or word is only needed once. Naturally, this option is also available at the phonological level.

Having identified the ∪ units as structural in nature, we may address the question of whether they fit in the pool of structural units that have been uncovered so far or whether they form a structural category of their own. The latter is the case. The slot level creates a sequence of ∪ units. The relationship among these units is entirely linear, no other form of organization, in particular no hierarchical one, is provided for. In contrast, the structural units discussed in the preceding section are quite different in kind. Take a VP as an example. When it is part of a larger structure (S), we have two structural levels with one structural unit subordinated to another. Such a multilevel organization is lacking at the slot level. We will therefore distinguish between two structural types—linear and multilevel structural units.

Phonemes are not the only fillers that require slots. In point of fact, all content units (with one exception) require them. We thus posit word slots, morpheme slots, and phoneme slots. However, there are no feature slots. This is for the aforementioned reason that features are not serialized in the production process. A more accurate description would consequently hold that all serializable content units require slots. This claim furnishes the explanation for why there are slots. As structural units, they are built up in the process of transforming a timeless lexical representation into a temporal representation that enables the sequential output of content units. Thus, slots as a set of lined-up positions are needed whenever the production process engages in serialization. It is worthwhile to add that this function should not be confused with the task of imposing a particular serial order, as is required for the generation of content units. In a sense, slots function as a prerequisite for serial-ordering processes, nothing more and nothing less.

This characterization makes slots an indispensable aspect of the language production process. Any model that purports to do without them should therefore be treated with scepticism. One such model was developed by Dell, Juliano, & Govindjee (1993) who assert that structural units are not necessary to account for certain speech error effects that have been regarded by many as being structurally motivated. Their model appears to successfully account for four error phenomena, but as the authors concede themselves, it is unable to generate the most important class of tongue slips, namely contextual errors. It is customary in speech error research to draw a distinction between contextual (syntagmatic) and noncontextual (paradigmatic) slips. Unlike the latter, the former are instigated by the context in which the error unit is embedded. Berg (2003a) reports for his German data that as many as 96.3% of all phonological slips, the error category that is at the heart of Dell et al.'s article, are contextually determined, whereas only 3.7% are not. This means that the Dell et al. model is seriously undermined by its inability to generate 26 out of 27 errors! Its failure to deal with contextual slips is certainly not coincidental. As argued earlier, the job of slots is to support a sequential representation. If no provision is made for slots, a model may produce nonsequential errors but must fail on the sequential ones. Precisely this happened in Dell et al.'s case.

Another model that manages without the slot-filler distinction is that of Vousden, Brown, & Harley (2000). It uses an oscillatory mechanism to control the serial order of phonemes and replaces the linear slots (or equivalently, frames) with what we may call temporal slots. The selection of phonemes occurs at certain prespecified moments in time. This model successfully accounts for a number of speech error effects including the parallel syllable structure constraint whereby segments preferentially interact with segments from similar structurally defined positions (see section 5.4.1.1 and beyond). Vousden et al. assume that structurally similar frames are selected at structurally similar times (metaphorically speaking) for example, at each full hour.

There are at least three empirical effects that prove difficult to capture. One is that it is unclear how this model handles addition and omission errors.[8] These word-shape errors inevitably alter the slot structure of the target word. Take the case of an addition error. It cannot help but desynchronize the system in that a segment that is intended to be produced at, let us say, 7 o'clock (e.g., a syllable-initial consonant), will be produced at 7:20 if it occurs after the addition error. That is, it would no longer be syllable-initial according to the logic of the model. However, the parallel syllable structure constraint is not banned from addition errors because the production system selects a different frame in the case of an addition error. It is hard to see how a frameless model can produce the requisite flexibility for this output variability.

The second problem is also one of lack of flexibility. As Berg & Abd-El-Jawad (1996) showed, languages may differ quite radically in the extent to which they obey the parallel syllable structure constraint. Unfortunately, Vousden et al. made no attempt to incorporate this variability into their model and in fact, it is hard to see how it could be incorporated, given that the lapsing of time is conceived of as a structured process. By implication, different notions of time would have to be invoked for different languages, which is hardly an attractive solution.

Finally, as will be argued immediately in the following discussion, slots are not neutral entities but may code certain types of information. This, of course, is utterly impossible in a model without slots. In view of these problems, it may be concluded that the Vousden et al. model does not seem to be able to supersede the slots-and-fillers approach, even though it has certain appealing properties.

So far, slots have been characterized as linearly arranged place holders that accommodate content units. The next issue is whether slots (at a particular level) accept all fillers or whether they are more "choosy." Choosiness is ordinarily conceived of as a specification on the slot that restricts the possible interactions between slots and fillers. Basically, three types of specification are conceivable. Slots may be completely unspecified, minimally specified, or maximally specified. An unspecified slot obviously accepts any filler whatsoever, a minimally specified slot imposes coarse-grained restrictions, and a maximally specified slot fine-grained restrictions on the nature of acceptable fillers. It is notable that the middle position has the most to recommend it. Unspecified slots make the prediction that, contrary to fact, anything goes. A prefix, for example, does not substitute for a stem in slips of the tongue. What we do find is that prefixes are replaced with other prefixes. This is a quite strong indication that this slot is geared to prefixes and nothing but. Information pertaining to the general class of an item may be viewed as a minimal specification. A maximal specification not only restricts the set of possible fillers to prefixes in general but to particular prefixes, for example, all disyllabic or reversative ones. Evidence for such maximal specification has been conspicuously lacking, whereas evidence for minimal specification is quite strong. In addition to the like-with-like constraint, some aspects of errors cannot be easily explained without taking recourse to slots. This can be illustrated on the basis of (28), a tongue slip from German.

(28) Welche *Erwartung*— welche Reaktion hast Du denn erwartet?
which expectation which reaction have you then expected
'Which reaction did you expect?'

Number (28) involves the anticipation of the lexeme *erwart(en)* 'to expect'. The remarkable feature of this slip is the occurrence of the suffix *–ung* in the error word *Erwartung* 'expectation,' which cannot be motivated by the source element because the latter has a verbal suffix. However, *-ung* is a nominal suffix. How else can we account for the emergence of *-ung* unless by assuming that it was retrieved on the basis of the information that a noun is to be produced? This information can only be extracted from the target slot for the noun *Reaktion* 'reaction.' This slot must therefore be specified for nominal.

Parallel cases occur in phonology. The boundary between consonants and vowels may occasionally be crossed, as in (29).

(29) Berkeley b*r*us. for: Berkeley bus. (from Stemberger, 1983b)

The curious aspect about (29) is the appearance of the postinitial rhotic in *brus*. Stemberger (1983b) argues that it stems from the syllabic /r/ in *Berkeley*. In this view, a phoneme that was associated with a V slot perseverated into a newly created C slot. This re-association induced a change in the phonetic nature of the misordered segment. Specifically, it took on a more consonantal quality. This can only be explained if the postinitial slot is specified as consonantal.

Comparable cases have not been found in the area of morphology. In all likelihood, the reason for this lack is the absence of a link between the different classes of units. At the lexical level, a link is created between nouns and verbs in virtue of their using the same morphological base (compare *erwart+en* and *Erwart+ung* in [28]). At the phonological level, the versatility of some segments forges a link between consonants and vowels (witness the rhotic in [29]). At the morphological level, however, there does not seem to be any such factor that could establish some common ground between prefixes and suffixes. As a consequence, interactions between them have not been observed.[9]

As an interim summary, it may be conjectured that slots are specified for the following major-class features, depending on the linguistic level they belong to.

(30)	Word slots:	noun, verb, adjective, etc.
	Morpheme slots:	prefix, suffix, stem
	Phoneme slots:	consonant, vowel.

In addition to these general properties, slots may code somewhat more specific information such as person on verbs, case on pronouns, and number on nouns. The logic of the argument is highly similar to the one made in connection with the major-class features. Certain aspects of tongue slips cannot be explained unless the malfunction is allowed access to information that is not inherent in the misordered unit but imposed on it by the slot in which it appears. The following example is again from German because this language has elaborate enough morphophonological paradigms to bring home the following point.

(31) daß sie die Augen*schüsse*— die Augenzeugen einfach niedergeschossen haben.
that they the eye-shots the eye witnesses simply down-shot have
'that they simply shot (down) the eye witnesses.'

Case (31) is not unlike (28) in that a verbal source ends up as a nominal intruder. What is particularly intriguing about (31) is the occurrence of the <ü> (= /y/) in the error word *Augenschüsse*. Whereas the verbal paradigm of the misordered element knows only /i:/ and /ɔ/ as stem vowels, the nominal paradigm knows /u/ and /y/. The claim that the slot of the error word was specified for nominal thus accounts for the vowel change as such, but it fails to arbitrate between /u/ and /y/. Note that the back vowel is appropriate for the singular (*Schuß* 'shot') whereas the front vowel is appropriate for the plural (*Schüsse* 'shots'). The explanation for the appearance of the /y/ lies in the hypothesis that this slot is not only specified for nominal but also for plural.

An exhaustive examination of the types of information for which slots may be specified would detract us from the main line of reasoning. Suffice it to say that slots may be specified for both major-class features and grammatical categories such as case and number. However, slot specification does not appear to go any further.

The function of slot specification is relatively easy to ascertain. It serves to facilitate the association process between slots and fillers. Slot specification considerably narrows down the range of potential candidates for selection. For example, knowing that a verb is needed for a particular slot eliminates no less than four fifths of the entire vocabulary. Hence, a minimum slot specification reaps good benefit. However, would this not imply that a maximum slot specification would be maximally efficient to the point of making the selection process immune to error? Although this may be so, the real question is not which level of detail should be, but *can* be attained. On account of their generality, the major-class features are the easiest to access and therefore available at an early point in processing to specify the slots yet to be filled (Berg, 1992a). More detailed pieces of information would become available only later when a lexical unit has been located. But then there would be no point in specifying a slot for an element that has already been retrieved.

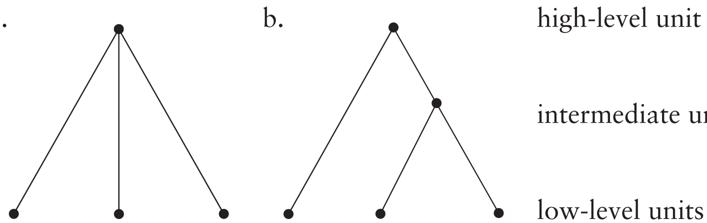
The underlying assumption here is that slots derive their information from the content units that they accommodate. Slot specification for major-class features is possible because this information is available prior to the more specific information and can thus be used to aid the retrieval of particular items. It may well be that a similar mechanism accounts for the other types of slot specification mentioned earlier. However, it is doubtful that all these types are processed alike because they fall on a continuum from more lexical to more syntactic. The general rule appears to be that the greater the syntacticity of an information type, the greater the likelihood of its acting as a slot specifier. For instance, as case is more syntactic than number on nouns, one might expect more case than number accommodations.

Summing up, a case has been made for distinguishing between two types of structural units—multilevel and linear ones. The focus in this section has been on the latter. Slots are structural units because they are not part of the long-term memory representation but are constructed in the process of language production. They have been argued to be an essential part of the serialization process even though they are not responsible for generating a certain serial order. In the production process, they are specified for major-class features and for grammatical-category information. Slot specification is understood here as alleviating the problem of lexical retrieval. In the following, the multilevel units will be paid closer attention.

1.5 THE WHYS AND WHEREFORES OF MULTILEVEL STRUCTURAL UNITS

At first sight, the reason that structure is required appears obvious enough. As argued previously, the function of structure is the gluing together of small units to form larger ones. An illustrative analogy might be the concrete that is used to build houses from bricks. This function seems evidently true in the case of syntax, given the fact that words are essentially the largest long-term memory units and that, by implication, sentences have to be created *de novo*. However, this cannot be the whole story. If it were, structure would only be needed above the word level. It would be superfluous below the word level because it makes no sense to create smaller structures (e.g., rimes) from larger content elements (e.g., monosyllabic words). Why, then, does a simple model consisting of content units but lacking structural units not suffice at the phonological level? The answer is that even in the face of important differences between syntax and phonology, both levels have very similar problems to solve. As all content units (save features) have to be serialized, the processor faces the same task in linearizing the phonemes in a word as in linearizing the words in a sentence. This is so irrespective of where the information that is fed into the serialization mechanism comes from. As argued in the preceding section, the need for serialization is the *raison d'être* of slots. Serialization thus explains the existence of phoneme slots, though not the existence of multilevel structural units. To appreciate the function of the latter, it is expedient to compare the two diagrams depicted in (32).

(32) a.



What is the advantage in having an intermediate node rather than none? Note at the outset that this node is *not* necessary for generating the correct serial order of the low-level units. Both the more complex structure (32b) and the less complex structure (32a) are suited to this task. The fundamental difference between (32a) and (32b) is the difference in advance planning that they afford of the linguistic material to be outputted. The flat structure affords no advance planning at all. The three low-level units are processed in strict succession, with no temporal overlap between them. When the first unit is selected, the production system leaves upcoming elements unattended to. The same is true of the second unit. All the system does after the selection of the preceding unit is pick the next unit to be produced until the final element has been reached. Thus, the representation in (32a) affords no possibility of looking ahead. It should be emphasized that although this is a disadvantage in certain respects (to be specified later in the book), it by no means disables the production system.

Representation (32b) has different implications. Its major strength is that it supports parallel processing. While the first unit is being processed, the system can concurrently access the intermediate node. No interference is expected between these two simultaneous processes because the one is directed at selecting a terminal node for output, whereas the other is directed at raising the activation level of a structural node. The great advantage of having an intermediate node available is that it provides the system with information about the number and (partly) the nature of the imminent elements. Assuming that an intermediate node is always branching, the system knows that at least two low-level units are in the offing and that, in the case of a rime node, the first is a vowel and the last a consonant. This knowledge allows the system to plan ahead reliably. In this way, it gains access to information about upcoming material at a very early moment in time. This information may be beneficial when decisions have to be quickly made that cannot be based on local considerations alone.

Three phenomena may suffice to illuminate this point. First, knowledge of the phonological make-up of the upcoming word serves as input to the choice of the appropriate allomorph. For example, the English indefinite article has the two variants *a* and *an*, the selection of which depends on whether the following word begins with a consonant or a vowel. This is an instance of advance planning that requires little look-ahead but the availability of fairly specific phonological information. The second example concerns rhythm. In a stress-timed language like English, knowledge of the number and stress value of the upcoming syllables is necessary for programming the duration of each syllable (see e.g., Nooteboom, 1995). This task requires look-ahead at least up to the next stressed syllable as well as access to the phonological level even though the quantity tier need not be called up. The third example is intonation. To compute a smooth intonation contour, advance knowledge spanning a tone unit is indispensible. Because tone units can be quite long, this presupposes a good deal of syntactic planning. And because the intonation contour depends on the tonic syllable, some phonological processing is also required. As these three examples show, advance planning is necessary to cope with the exigencies of the linguistic system because it permits the processor to make well-reasoned decisions. This look-ahead capacity is the major reason for the emergence of multilevel structural units.

The metaphor that will henceforth be used to denote the degree of advance planning is the *planning window*. Structural units may be said to open the planning window.[10] The degree of opening varies with the hierarchical position of the structural unit. The higher the unit is in the hierarchy, the greater the size of the window. A corollary of this is that the lower the level of linguistic analysis, the smaller the window size and the more local the factors that influence a particular decision.

The planning window is intentionally defined in linguistic terms (i.e., in terms of the nature of the structural unit). Alternatively, it might be construed as a temporal notion and termed a time window. These two notions correlate to a certain extent. A larger unit takes more time to be produced. A larger structural unit is therefore capable of opening the window for a longer time. However, the correlation is not perfect as the length of time for which a given unit is available may vary (depending on speech rate for example). Thus, the concept of a temporal window is more difficult to handle than that of a planning window. In view of the emphasis on structure, the idea of a planning window proves to be more consonant with the aims of this study and will therefore be preferred.

What light does this account shed on the assumed function of structural elements to create units that exceed the size of those that are stored in long-term memory? A somewhat surprising implication of the preceding analysis is that this function is of limited importance. If it were the main reason for the emergence of structure, we would not expect to find it at the phonological level. The observation that structure manifests itself in exactly the same way in syntax and phonology allows us to argue that the function of increasing the size of linguistic units is something of a side effect of planning the linearization of language. Of course, this side effect is highly welcome and even essential for the creative use of language, but the planning process would be basically the same if sentences were stored in long-term memory and retrieved in the same way as words.

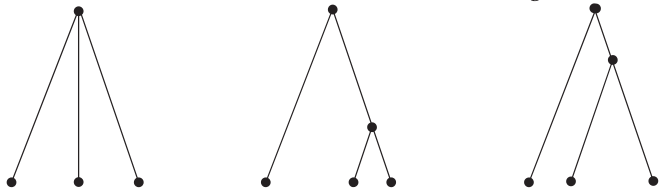
1.6 IMPLEMENTING STRUCTURE

In the foregoing, the building blocks of language and the function they perform have been examined. The next step is to probe into the mechanisms by which multilevel structural units are implemented. A spreading activation framework (e.g., Anderson, 1983) is optimally suited to this purpose. Returning to the diagrams in (32), we may ask what influence an intermediate node has on the activation spread in a network. There are basically three effects. The first is that an intermediate node slows down the activation process. Activation has to be built up on a node and this process takes time (Ratcliff & McKoon, 1981). As a result, the second and third low-level nodes are activated later than the first. The second effect is what may be called the equalizing power of nodes. A(n intermediate) node simultaneously passes on essentially the same quanta of energy to its subordinate nodes. Third, the further down we move in the linguistic hierarchy, the greater the degree of coactivation of units that are dominated by the same node. This is mainly a frequency effect because lower units are used more often than higher units (given their part–whole relationship). Hence, the degree of coactivation of the second and third low-level nodes is higher in (32b) than in (32a).

This notion of coactivation is of crucial importance in that it determines the cohesiveness of linguistic units (Berg, 1989b). The higher their degree of coactivation, the stronger their cohesiveness. This principle follows from the way linguistic units are selected. Selection is generally held to be a function of activation levels. The unit that is most strongly activated at a certain point in time is selected for production (e.g., MacKay, 1987). When two units are strongly activated at the same time and slots are available for each of them, they stand a good chance of being selected together. This, of course, is what has been referred to as cohesiveness in the earlier analysis of slips of the tongue.

A real asset of the spreading activation framework is its ability to deal with gradience, and this is particularly true in the case of cohesiveness. As shown in section 1.3, linguistic units display widely differing degrees of cohesion. These can be nicely modelled by postulating variable activation levels for structural units. An intermediate node as in (32b) with a relatively low activation level engenders a low degree of cohesiveness of its subordinate nodes (and vice versa). Thus, it is the activation level of structural units that determines the cohesiveness of sequences of content units.[11] This gradualness may be graphically represented by the "snapshots" in (33). Cohesiveness is expressed on paper by the vertical position of the intermediate node between the superordinate and subordinate nodes. The higher its position, the more activation it is assumed to have.

(33) a. no cohesiveness b. low cohesiveness c. high cohesiveness



There are two principal factors that influence the activation levels of structural units. The first was introduced in section 1.3. Because activation takes time to build up, a later stage in the production process will see higher activation levels (and hence more cohesiveness) than an earlier one will. In a word, more time, more structure, more cohesiveness.[12] Furthermore, the level of activation may be a function of linguistic orthodoxy. Basically, this concept reflects the frequency with which content units are put together to form larger structures. For example, the plural–singular ratio of *children–child* is higher than that of *snows–snow*. The plural form *children* may therefore be assumed to reach a higher activation level than the plural form *snows*. This difference brings us back to the two slips of the tongue (8) and (9), which are reproduced here as (34) and (35) for convenience.

(34) Well you can cut *rain* in the *trees*. for: trees in the rain. (from Garrett, 1982)

(35) He doesn't have any *closet*s in his *skeleton*. for: skeletons in his closet. (from Stemberger, 1985)

Why is the plural word cohesive in (34) but incohesive in (35)? For one thing, *trees* is a "good" plural word, for another *rains* is a "bad" plural word. Both these arguments conspire to impart cohesiveness to a structure (i.e., an inflected word) that is typically not cohesive (see section 1.3). The default case in (35) therefore needs less motivation than the untypical case (34). In fact, the two interacting nouns in (34) are less susceptible to pluralization than *trees* but do not repel pluralization as does *rain*. Thus, the reason for cohesiveness in (35) and lack thereof in (34) is the linguistic orthodoxy of the outcome. If cohesiveness leads to a "better" error, it may outweigh the more common alternative of breaking up inflected words.

There is no need to assume a competitive relationship between the time hypothesis and the orthodoxy hypothesis. Both influence activation levels in the same way—more time, more activation, and more orthodoxy, more activation. Both are of similar generality. The time hypothesis holds that any activation process, in particular in the case of structural units, takes time. The orthodoxy hypothesis claims that a more frequent unit will amass more activation in the same time than a less frequent one. Many of the slips of the tongue are compatible with both hypotheses. The errors in (34) and (35) can be explained by the orthodoxy hypothesis as was done previously as well as by the time hypothesis by simply assuming that (34) occurred at a later temporal stage than (35). Whether the two hypotheses jointly produce the difference between (34) and (35) or whether one of them plays a larger role in this process than the other is an open issue. In any event, the main point of the preceding discussion is unaffected by the ultimate answer. The cohesiveness of two content units is determined by their degree of coactivation, which in turn is determined by the activation level of the superordinate structural unit.

1.7 REVIEW AND PREVIEW

The structural model of language production that has been sketched out in this introductory chapter is firmly rooted in psychology. It stands or falls on the assumption that not all units that are standardly considered relevant in linguistics are part of a speaker's long-term store. As a consequence, some units have to be created during the preparation to speak whereas others can be retrieved ready-made from long-term memory. This fundamental difference underlies the distinction between content and structural units. The latter can be subdivided into slots and multilevel elements. Slots are reserved for serializable content units and are linearly concatenated. Their task is to enable the linearization of language. Multilevel structural units function to increase the planning span for an utterance.[13] The larger these structural units are, the further ahead the speaker can plan. Thus, the present model recognizes the distinction between hierarchical and linear representations that is variously made in both the linguistic and the psycholinguistic literature (Martin, 1972; Falk, 1983; Vigliocco & Nicol, 1998; Kathol, 1999; Hartsuiker, Kolk, & Huiskamp, 1999). A noteworthy property of the model is its gradience (see also Bolinger, 1961 from the linguistic and Yantis & Meyer, 1988 from the psychological perspective).[14] Because of the model's reliance on the activation metaphor, structural units are not either present or absent but more or less strongly activated. The variable levels of activation entail differences in the model's output.

The processing roots of the model imply that all linguistic units posited are claimed to be psychologically real. This is particularly true of structural nodes. In fact, Bock & Loebell (1990) were able to show that (syntactic) constituent structure can be primed by constituent structure, thereby demonstrating its psychological reality.

The relationship between content and structural units is depicted in Figure 1.1 (see opposite page).

The various aspects of information flow in Figure 1.1 may be depicted thus. First, the vertical connections between the content units make sure that all serializable elements are correctly "interpreted." If, let us say, the /t/ cannot be used to access the feature [alveolar], the production system will degrade into muteness. Second, the content units are responsible for the creation of slots. The activation flow from fillers to slots is a natural consequence of the claim that content units are stored in long-term memory, whereas structural units are not. Clearly, the former can assist in the creation of the latter, though not vice versa. Once a slot has been created it can be specified, and thereby constrain lexical, morphological, and phonological access. Third, the slots are the terminal elements on the basis of which multilevel structural units can be erected in bottom-up fashion. Finally, there might be direct links between the structural units from different levels (not depicted in Figure 1.1). Although such links are not necessitated by any theoretical argument, empirical data in support of them will be adduced in