Word Grammar

New Perspectives on a Theory of Language Structure

edited by Kensei Sugayama and Richard Hudson

continuum

Continuum
The Tower Building
11 York Road
London SE1 7NX

15 East 26th Street
New York
NY 10010

First published 2006

British Library Cataloguing-in-Publication Data
A catalogue record for this book is available from the British Library.

ISBN: 0-8264-8645-2 (hardback)

Library of Congress Cataloguing-in-Publication Data
To come

Typeset by BookEns Ltd, Royston, Herts.
Printed and bound in Great Britain by MPG Books Ltd, Bodmin, Cornwall

The problem of the word has worried general linguists for the best part of a century.
—P.H. Matthews

Contents

Trask, Robert L. (1993), *A Dictionary of Grammatical Terms in Linguistics*. London: Routledge.
Trier, Jost (1931), *Der Deutsche Wortschatz im Sinnbezirk des Verstandes. Von den Anfängen bis zum 13. Jahrhundert*. Heidelberg: Winter.
Weisgerber, Leo (1927), 'Die Bedeutungslehre - ein Irrweg der Sprachwissenschaft'. *Germanisch-Romanische Monatsschrift*, 15, 161–83.
Williams, Edwin (1991), 'Meaning categories of NPs and Ss'. *Linguistic Inquiry*, 22, 584–7.

Notes

1 The SSP given here turns out not to be able to account for all cases of linking. It is revised in Holmes (2005: 44).
2 Note that in Biber *et al.* the VP category subsumes the 'verbal complex' (main verb and any auxiliaries), but not any complements or other postdependents of the verb.
3 Of course, if a particular speaker knows the words ACTOR and THEME (as metalinguistic terms), then they must have these relationships in their lexicon, since they are (or should be!) the meanings of the relevant terms.

6 Word Grammar and Syntactic Code-Mixing Research

EVA EPPLER

Abstract

This chapter aims to show that WG is preferential over other linguistic theories for the study of bilingual speech. Constituent-based models have difficulties accounting for intra-sentential code-mixing because the notions of government and functional categories are too powerful and rule out naturally occurring examples. Properties of WG which make this syntactic theory particularly well suited for code-mixing research are the central role of the word, the dependency analysis, and several consequences of the view of language as a network which is integrated with the rest of cognition. A qualitative and quantitative analysis of *because* and *weil* clauses shows that code-mixing patterns can be studied productively in WG.

1. Introduction

Intra-sententially CODE-MIXED data, i.e. utterances constructed from words from more than one language, pose an interesting problem for syntactic research as two grammars interact in one utterance. Based on a German/English bilingual corpus,[1] I will show in section 2 of this chapter that constraints on code-switching formulated within Phrase Structure Grammar frameworks (Government and Binding, Principles and Parameters, Minimalism) are too restrictive in that they rule out naturally occurring examples of mixing.

In section 3 I will discuss aspects of WG that make it particularly well suited for the syntactic analysis of intra-sententially mixed data. WG facilitates the full syntactic analysis of sizeable corpora and allows us to formulate hypotheses on code-switching which can subsequently be tested on data. All findings are supported by quantitative data.

As the word order contrast between German and English is most marked in subordinate clauses, I focus on examples of this construction type in section 4. I will show that code-mixing patterns can be studied productively in terms of WG: WG rules determining the word order in German/English mixed clauses hold in relation to my corpus and are supported by evidence from other corpora. The main section of this chapter focuses on *because* and *weil* clauses. A comparison of the mixed and monolingual clauses reveals that German/English bilinguals who engage in code-mixing recognize and utilize structural

congruence at the syntax-pragmatics interface. They predominantly mix in a construction type in which the word order contrast between German (SOV) and English (SVO) is neutralized.

2. Constituent Structure Grammar Approaches to Intra-Sentential Code-Mixing

The question underlying grammatical code-switching research is whether there are syntactic constraints on code-mixing. Some of the hypotheses on intra-sentential code-switching have been formulated in informal frameworks of traditional grammatical notions; others are derived from assumptions underlying specific modern syntactic theories. In this section I will review the main phrase structure grammar approaches to code-mixing and show that the constraints formulated within them do not account for the data.

DiSciullo, Muysken and Singh (1986) propose to constrain code-switching by government, the traditional assumption behind X-bar theory. They initially used the Chomsky (1981: 164) formulation of government 'α governs γ in $[\beta \ldots \gamma \ldots \alpha \ldots \gamma \ldots]$, where $\alpha = X^0$, and α and γ are part of the same maximal projection'. The X-bar assumption that syntactic constituents are endocentric is important for the formulation and working of the government constraint. Heads not only project their syntactic features onto the constituent they govern, but also their language index. The language index is assumed to be something specified in the lexicon (DiSciullo *et al.* 1986: 6), since the lexicon is a language-specific collection of elements. For code-switching purposes the Government Constraint was formalized in DiSciullo *et al.* (1986: 6) as $[X^p\ Y^p]$, where X governs Y, and p and q are language indices. The nodes in a tree must dominate elements drawn from the same language when there is a government relation holding between them.

The Government Constraint predicts that ungoverned elements, such as discourse markers, tags, exclamations, interjections and many adverbs, can easily be switched. This prediction is also supported by my data (see also Eppler 1999) and most other bilingual corpora. However, the Government Constraint also predicts that switches between verbs and their objects and/or clausal complements, and switches between prepositions and their NP complements, are ungrammatical. Examples violating these predictions from my corpus are:

(1)	*TRU:	so /|/ so **you have**		eine	Übersicht.	Jen2.cha, line 133
				an	overview	
(2)	*DOR:	**I wonder,**	wem	sie	nachgradt.	Jen2.cha, line 1531
			whom	she	takes after	

or in the other direction, i.e. from a German verb to an English clausal complement:

(3)	*MEL:	ich hab(e) gedacht **there is going to be a fight.**		Jen1.cha, line 987
		I have thought		

(4)	*TRU:	der	hat	über **faith + healing**	gesprochen.	Jen2.cha, line 2383
		he	has	about	spoken	

The original inclusion of functional categories in the class of governors ruled out code-switches which are also documented in my data, e.g. between complementizers and clauses that depend on them, as in (5):

(5)	TRU:	**to buy yourself in means that +...**					Jen1.cha, lines 977ff
	DOR:	du	kannst	dich		nochmal	einkaufen.
		you	can	yourself		once more	buy in

and the domain of government was too large. The above formulation of the government constraint includes the whole maximal projection and thus, for example, bans switching between verbs and location adverbs, again contrary to the evidence. Therefore a limited definition of government, involving only the immediate domain of the lexical head, including its complements but not its modifiers/adjuncts, was adopted and the Government Constraint was rephrased (Muysken 1989) in terms of L-marking:

*[Xp Yq], where X L-marks Y, and p and q are language indices

Muysken and collaborators thus shifted from an early and quite general definition of government to the more limited definition of L-marking in their formulation of the Government Constraint. L-marking restricts government to the relation between a lexical head and its immediate complements. Even the modified version of the government constraint in terms of L-marking is empirically not borne out, as we see from the following example:

(6)	TRU:	das	ist	**painful.**	Jen3.cha, line 1879
		this	is		

Muysken (2000: 25) identifies two main reasons why the Government Constraint, even in its revised form, is inadequate. The main reason is that CATEGORIAL EQUIVALENCE[2] undoes the effect of the government restriction. The Government Constraint is furthermore assumed to insufficiently acknowledge the crucial role functional categories are supposed to play in code-mixing.

Functional categories feature prominently in several approaches to code-mixing. Joshi, for example, proposes that 'Closed class items (e.g. determiners, quantifiers, prepositions, possessive, Aux, Tense, helping verbs, etc.) cannot be switched' (1985: 194). Myers-Scotton and Jake (1995: 983) assume that in mixed constituents, all SYSTEM MORPHEMES[3] that have grammatical relations external to their head constituent (i.e. participate in the sentence's thematic grid) will come from the language that sets the grammatical frame in the unit of analysis (CP). And Belazi, Rubin and Toribio (1994) propose the Functional Head Constraint.

Their model is embedded in the principles and parameters approach. Belazi, Rubin and Toribio (1994) propose to restrict code-mixing by the

feature-checking process of f-selection. In Belazi, Rubin and Toribio's model, language is a feature[4] of FUNCTIONAL heads that needs checking like all other features. The Functional Head Constraint (Belazi, Rubin and Toribio (1994: 228)) is formulated as follows:

> The language feature of the complement F-selected by a functional head, like all other relevant features, must match the corresponding feature of that functional head.

Code switching between a lexical head and its complement proceeds unimpeded in this model.

Because many inflectional morphemes were treated as independent functional heads in the principles and parameters approach, Belazi, Rubin and Toribio (1994) subsume the FREE MORPHEME CONSTRAINT[5] (Sankoff and Poplack 1981) under their functional head constraint: switching is disallowed between an inflectional morpheme and a word-stem. A counter-example to this restriction from my corpus would be:

(7) *DOR: wir **suffer-n** da alle. Jen2.cha, line 904
we suffer INFL MP all

Like all researchers working on Spanish/English and Arabic/French code-mixing, Belazi, Rubin and Toribio (1994) have to deal with the different placement of adjectives pre- or post-modifying nouns in the language pairs they are working on. Their data indicate that switching is possible when the adjectives and nouns obey the grammars of the languages from which they are drawn. This leads them to supplement the Functional Head Constraint with the WORD-GRAMMAR INTEGRITY COROLLARY,[6] which states that 'a word of language X, with grammar G, must obey grammar G' (Belazi, Rubin and Toribio 1994: 232).

Like the Government Constraint, the Functional Head Constraint rules out switches between complementizers and their clausal complements. Therefore example (5) provides counter-evidence to this constraint. It also rules out switches between infinitival *to* and its verbal complement, examples of which are also attested in my corpus.

(8) *LIL: **you don't need to** wegwerfen. Jen2.cha, line 2555
throw away

The Functional Head Constraint furthermore rules out switches between determiners (including quantifiers and numerals) and nouns. As nouns are the most frequently borrowed or switched word class, counterexamples abound in my and many other corpora.

MacSwan (1999: 188), working within the minimalist framework, also assumes that code-switching within a PF component is not possible. This PF Disjunction Theorem amounts to the same effect as the Free Morpheme Constraint (Sankoff and Poplack 1981) and the various restrictions on switching

between stems and morphologically bound inflectional material. Examples (7) and (9) are therefore clear violations of the PF Disjunction Theorem.

(9) *DOR: sie haben einfach nicht ge**#bother-ed**. Ibron.cha, lines 1012, 14
they have simply not

The minimalist framework he is working in forces MacSwan (1999) to preserve constituent structure, but he acknowledges the advantages of a system of lexicalized parameters for the analysis of code-switching.

In this section I reviewed approaches to code-mixed data that crucially depend on constituency structure/maximal projections (DiSciullo, Muysken and Singh 1986) and functional categories (Belazi, Rubin and Toribio 1994). I showed that these constraints and models are too restrictive in that they rule out naturally occurring examples of intra-sentential code-mixing. The 'government constraints' (DiSciullo, Muysken and Singh 1986; Muysken 1989) were found to be too restrictive when tested against natural language data because the government domain was too large. Models, approaches and constraints based on functional categories (Joshi 1985; Myers-Scotton 1993; Belazi, Rubin and Toribio 1994) fall short of accounting for the data available and are unsatisfactory because none of the definitions of functional categories that have been offered (in terms of function words, closed class items, system morphemes or non-thematicity) work. They either define fuzzy categories where a sharp distinction would be needed, or they conflict with the data. Complementizers and determiners, the two most commonly quoted examples of functional categories, provide most of the counterexamples.

For these reasons a syntactic theory that rejects constituency structure and does not recognize functional categories (Hudson 2000) seems an interesting and promising option to explore. In the next section I will review other aspects/characteristics of WG which are perceived to make this theory of sentence structure more suitable for the analysis of (monolingual and) code-mixed data than other theories.

3. A Word Grammar Approach to Code-Mixing

The main reason why I chose WG for the syntactic analysis of my data is because this theory of sentence structure takes the word as a central unit of analysis. In WG, syntactic structures are analysed in terms of dependency relations between single words,[7] a parent and a dependent. Phrases are defined by dependency structures which consist of a word plus the phrases rooted in any of its dependents. In other words, WG syntax does not use phrase structure in describing sentence structure, because everything that needs to be said about sentence structure can be formulated in terms of dependencies between single words. For intra-sententially switched data this is seen as an advantage over other syntactic theories because each parent only determines the properties of its immediate dependent. Language specific requirements are thus satisfied, if the particular pair of words, i.e. the parent and the dependent, satisfy them. A word's requirements do not project to larger units like maximal projections/

phrasal constituents. If we want to formulate constraints on code-switching within WG, they have to be formulated for individual types of dependency relations. Because they do not affect larger units, they are less likely to be too restrictive than constraints affecting whole phrasal constituents. One of the main problems of constituency based models, i.e. over-generalization through phenomena like government chains, therefore cannot occur in a WG approach to code-mixing.

The central role of the word in WG moreover means that words are not only the largest units of WG syntax, but also the smallest. In contrast with Chomskyan linguistics, syntactic structures do not, and cannot, separate stems and inflections. Furthermore, at least as far as overt words are concerned, WG rejects the notion of functional category. Hudson (2000) shows that this notion is problematic, because it has never been defined coherently and because all the individual categories that have been given as examples (e.g. complementizers) raise serious problems. For the same reasons, constraints on intra-sentential code-switching based on functional categories (Joshi 1985; Belazi, Rubin and Toribios 1994) and models of code-switching that crucially depend on the distinction between system and content morphemes (Myers-Scotton 1993) run into serious empirical difficulties (see section 2). Because WG is an example of 'morphology-free syntax' (Zwicky 1992: 354) which rejects the notion of functional categories, a WG approach to intra-sentential code-switching cannot over-emphasize the role of inflectional morphemes.

Words being the only and central unit of analysis in Word Grammar furthermore benefits code-mixing research in a purely pragmatic way. The majority of research in this area is based on sizable natural language corpora. Because the only units that need to be processed in WG are individual words and larger units are built by dependency relations between two words which can be looked at individually, a WG approach to intra-sentential code-mixing requires less analysis than constituency-based modes. This facilitates the analysis of large-ish corpora. Eppler (2004), for example, is based on a WG analysis of a 22,000 word corpus.

For intra-sententially mixed sentences a dependency analysis is furthermore seen as an advantage over phrase structure grammar frameworks because it highlights the functional relations between words (from the same or different languages) rather than code-switch points. Constituency-based models describe and/or constrain intra-sentential code-switching by disallowing switches between, for example, PP and NP (see section 2). A WG analysis would note a switched complement relation which is grammatical, if the preposition and the determiner/(pro-)noun involved in it satisfy the constraints imposed on them by their own language. To start to understand what is going on in intra-sentential code-switching, it seems more beneficial to gain an insight into which syntactic relations are frequently or rarely switched, rather than to increase our knowledge about points in sentences where switching does not occur.

Another characteristic of WG is that dependency analyses have a totally flat structure. A single, completely surface structure analysis (with extra dependencies being drawn below the sentence-words) is seen as benefiting

WG over other theories of language structure for code-mixing research: linguists working on code-mixing during times when Chomskyan frameworks still stressed the difference between surface and deep structure did not know what to do with D-structure, because code-switching clearly seems to be a surface structure phenomenon. Romaine (1989: 145) concludes her discussion of the government constraint with the statement 'data such as these [code-mixing data] have no bearing on abstract principles such as government [...] because code-switching sites are properties of S-structure, which are not base generated and therefore not determined by X-bar theory'. This problem does not emerge when one works with WG because of its totally flat, i.e. surface, analysis. A syntactic theory that shares properties of the linguistic phenomenon under investigation appears to be preferable to other syntactic theories; i.e. for a surface-structure phenomenon like code-mixing, a syntactic model that allows a single, completely surface analysis seems to be well suited.

Other aspects of WG which make this theory of sentence structure more suitable for the analysis of code-mixed data than other theories are derived from the WG view of language as a network which contains both the grammar and the lexicon and which integrates language with the rest of cognition.

This cognitive view of language as a labelled network has consequences for a controversial issue in psycholinguistic bilingualism research: the lexicons debate, i.e. whether bilinguals' lexical items/lemmas are stored in one or two lexicons. The network idea offers the advantage of viewing a bilingual's two languages as sub-networks, with denser links between lexical items from the same language and looser connections between lexical items from different languages.

This view of the bilingual lexicon(s) in combination with the multiple default inheritance system which WG operates on could possibly have enormous benefits for writing a psycholinguistically realistic grammar of a bilingual. The following exploration is just a sketchy idea as to how this could work and requires fleshing out, but the basic idea seems to work. Default inheritance allows us to build a maximally efficient system for bilinguals by locating the shared properties of words which 'belong' to different languages higher up the is-a hierarchy and the language specific properties lower down in this hierarchy. English *come* and German *kommen*, for example, are both verbs (is-a verb). They therefore share certain characteristics: they have a similar meaning ('move towards'), they both have tense (present or past), they have a subject and the subject tends to be a pre-dependent noun, etc. All these generalizable facts about German and English verbs can be located fairly high up in the is-a hierarchy. The features in which our two example words differ, for example that they have a different form (/kɒmən/ and /kʌm/ respectively), and that German *kommen*, when it is the complement of a subordinating conjunction or an auxiliary/modal, would be placed in clause final position, would be lower in the is-a hierarchy. Because of the way default inheritance works, characteristics of a general category are 'inherited' by instances of that category only if they are not overridden by a more specific (e.g. language-specific) characteristic. A fact located lower down in the inheritance hierarchy of entities or relations takes priority over one located above it. Thus we could maximize the bilingual system

by allowing generalization by default inheritance and ensure that the language specific properties would automatically override the general pattern. For bilinguals this system would have the advantage that the grammatical system of a Castilian/Catalan bilingual, for example, would have fewer overriding/blocking language specific properties listed than that of a German/English bilingual.[8]

WG furthermore aims at integrating all aspects of language into a single theory which is also compatible with what is known about general cognition; that is, language can be analyzed and explained in the same way as other kinds of knowledge or behaviour. For example, it is widely acknowledged that code-mixing is influenced by social and psychological factors (Muysken 2000) and a syntactic model that allows us to incorporate this kind of information is better suited to describe language contact phenomena than theories that deal exclusively with language. Knowledge of more than one language, and the use of more than one language in one sentence, can be analyzed and explained in the same way as knowledge of one language and monolingual language use. In other words, code-mixing is not seen as 'deviant'. Because WG aims to explain and analyze language in the same way as other kinds of social and psychological knowledge or behaviour, it is perceived to be more suitable for research into bilingualism than other models of syntax.

The WG view of language as a network of associations which is closely integrated with the rest of our knowledge lends itself particularly well to code-mixing research for another reason. It is a well accepted fact in this research paradigm that adult bilinguals know, first of all, which language the words they use belong to. Second, they know when to code-switch and when not to (code-switching as a MARKED or UNMARKED choice,[9] for example, Myers-Scotton and Jake 1995), or when they should be in MONOLINGUAL SPEECH MODE or when they can go into BILINGUAL MODE (Grosjean 1995). Third, bilinguals also know which mixing patterns are acceptable in their speech community and which are not (SMOOTH versus FLAGGED code-switching,[10] for example, Poplack and Meechan 1995). This knowledge about language use is obviously closely integrated with other types of (social) knowledge and a syntactic theory that views language as a part of the total associative network is clearly more suitable to explain these phenomena than other theories.

Viewing language as a sub-network (responsible for words) which is just a part of the total associative network creates another advantage of WG for the research paradigm under discussion in this chapter. This benefit is related to the fact that most code-mixing research is based on natural language corpora.[11] In contrast with most other kinds of grammar which generate only idealized utterances or sentences, WG grammar can generate representations of actual utterances. A WG analysis of an utterance is also a network; it is simply an extension of the permanent cognitive network in which the relevant word tokens comprise a fringe of temporary concepts attached by 'is-a' links; so the utterance network has just the same formal characteristics as the permanent network. This blurring of the boundary between grammar and utterance is quite controversial, but it follows from the cognitive orientation of WG. For work based on natural speech data it is seen as another crucial advantage of

WG over other theories which can only generate syntactic structures for sentences. From the examples quoted so far, it is obvious that the audio data this study is based on are transcribed as utterances, i.e. units of conversational structure. For the grammatical analysis, however, I assume that conversational speech consists of the instantiation of linguistic units, i.e. sentences. In other words, every conversational utterance is taken to be a token of a particular type of linguistic unit, the structural features of that unit being defined by the grammatical rules of either German or English. When using a WG approach to code-mixed data, one does not have to 'edit' the corpus prior to linguistic analysis. Any material that cannot be taken as a token of either a German or English word-form can be left in the texts, but if it cannot be linked to other elements in the utterance via a relationship of dependency, it is not included in the syntactic analysis. That is, all the words in a transcribed utterance that are related to other words by syntactic relationships constitute the sentences the grammatical analysis is based on. As far as I am aware, WG is the only syntactic theory that can (and wants to) generate representations of actual utterances, and facilitates the grammatical analysis of natural speech data without prior editing.

Another consequence of integrating utterances into the grammar is that a word token must be able to inherit from its type. Obviously the token must have the typical features of its type – it must belong to a lexeme and a word class, it must have a sense and a stem, and so on. But the implication goes in the other direction as well: the type may mention some of the token's characteristics that are normally excluded from grammar, such as characteristics of the speaker, the addressee and the situation. For example, we can say that the speaker is a German/English bilingual and so is the addressee; the situation thus allows code-mixing. This aspect of WG theory thus allows us to incorporate sociolinguistic information into the grammar, by indicating the kind of person who is a typical speaker or addressee, or the typical situation of use.

Treating utterances as part of the grammar has further effects which are important for the psycholinguistics of processing. The main point here is that WG accommodates deviant input because the link between tokens and types is guided by the 'Best Fit Principle' (Hudson 1990: 45ff): assume that the current token is-a the type that provides the best fit with everything that is known. The default inheritance process which this triggers allows known characteristics of the token to override those of the type. Let's take the deviant word /bʌsə/ in the following example:

(10a) *TRU: xxx **and** warum waren keine bus(s)e [%pho: bʌsə]? Jen3.cha, line 331
why were there no buses

/bʌsə/ is phonologically deviant for German (*Busse* is pronounced /busə/), and morphologically deviant for English, because the English plural suffix is *–(e)s*, not *–e*. Although this word is deviant,[12] it can is-a its type, just like any other exception. But it will be shown as a deviant example. There is no need for the analysis to crash because of an 'error'. The replies to *TRU's question clearly show that the conversation does not crash:

(10b) *LIL: xxx [>] wegen einer bombe.
*MEL: xxx [>] **a bomb scare.** Jen3.cha, lines 332-333

This is obviously a big advantage of WG for natural speech data.

Another characteristic of natural speech data – and code-mixed data in particular – is that they are inherently variant. Most syntactic theories aim at describing and explaining regularized and standardized linguistic data and therefore disregard inherent variability. Hudson (1997) outlines how a prototype-based network theory that is based on default inheritance and uses entrenchment, like WG, can incorporate variation.

One of the strengths of the network approach is that it allows links to have different 'strength'; these are an essential ingredient of the model of spreading activation, and are highly relevant to quantitative work. Hudson (1997) stipulates that a language user who observes variation will arrive at generalizations about this variation. Each part of a variable network structure has some degree of 'entrenchment' which reflects the experiences of the person concerned. The degree of entrenchment of a concept can be presented as a probability of that concept being preferred to any relevant alternatives. This is presented for word-final variable t/d loss in Figure 1, where the figures[13] in angled brackets present the probabilities.

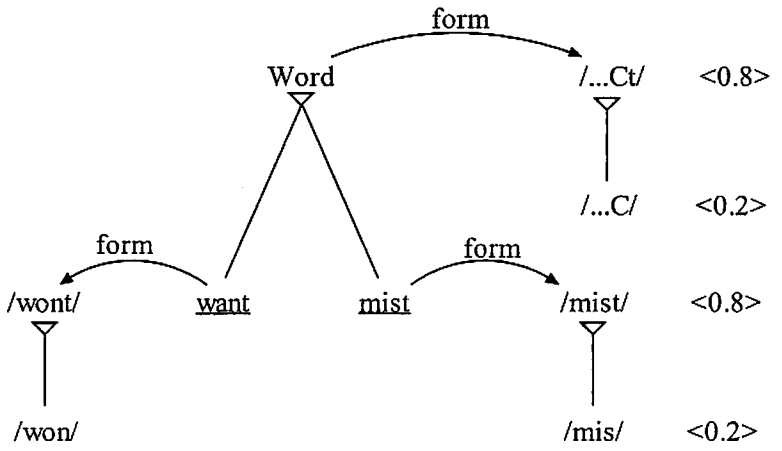

Figure 1 (Hudson 1997: Figure 5)

This analysis of variation is declarative and non-procedural and requires just two elementary operations: pattern-matching and default inheritance. Speakers and hearers need to know that alternative forms can be used instead of the basic form, and in a real life context the choice between them is influenced by the linguistic and social context. Figure 2 just hints at how these extra variables could be introduced.

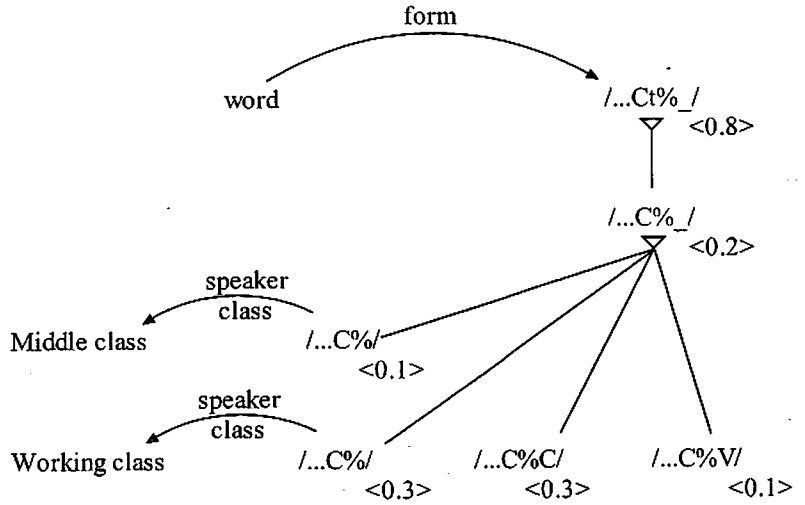

Figure 2 (Hudson 1997: Figure 6)

This model of inherent variability is possible because WG assumes that linguistic concepts are closely linked to non-linguistic concepts and carry quantitatively different entrenchment values. The reason why I find the proposed model so appealing is because it is a model of competence – not performance. Inherent variability is generally (rightly or wrongly) associated with performance, and to my knowledge there is no other model that presents variability and sociolinguistic information as part of a speaker's competence. I believe that linguistic variation that is influenced by social factors **is** part of every speaker's competence and a (more fleshed out) model of how speakers exploit their sociolinguistic competence is therefore required within linguistic theory.

In the following main section of this chapter I will present a quantitative/variationist and qualitative analysis of monolingual and code-mixed subordinate clauses. As none of the syntactic restrictions on code-switching proposed in the literature hold absolutely and universally, several recent studies in the field (Mahootian and Santorini 1996; MacSwan 1999; Eppler 1999) have reverted to the null hypothesis. I take the same approach. Formulated in WG terms, the null hypothesis assumes that each word in a switched dependency satisfies the constraints imposed on it by its own language.

Subordination was chosen as an area of investigation because the two languages in contact in this particular situation, German and English, display surface word order differences: English subordinate clauses are SVO whereas German subordinate clauses are SOV. The contrasting word order rules for English and German, stated in Word Grammar rules, are:

E1) In English any verb follows its subject but precedes all its other

dependents. This holds true for main as well as subordinate clauses and gives rise to SVO order in both clause types.

E2) Subordinators, e.g. *because*, require a following finite verb as their complement. A word's complement generally follows it.[14]

For German the most relevant rules[15] concerning word order in main and subordinate clauses are:

G1) A default finite verb follows one of its dependents but precedes all other dependents. This gives rise to a verb second (V2) word order in German main clauses.

G2) A finite verb selected by a lexical subordinator/complementizer takes all its non-verb dependents to the left, i.e. it is a 'late'[16] verb.

G3) Subordinators/complementizers, e.g. *daß*, select a 'late' finite verb as their complement.[17] According to G2 finite 'late' verbs follow all their non-verb dependents.

An example illustrating rules G1-G3 would be:

(11) Ich glaube nicht, daß wir die Dorit schon gekannt haben
I think not that weDorit already known have
Jen3.cha, line 83

The utterance initial main clause displays V2 word order. The finite auxiliary *haben* which depends on the subordinator *daß*, on the other hand, is in clause final position following all other constituents including non-finite verbs like *gekannt*. In English finite verbs in subordinate clauses do not behave differently from finite verbs in main clauses. Therefore we do not have to override the default rule E1 in the 'isa-hierarchy' of grammar rules. Because German finite verbs depending on a subordinator take a different word order position than 'independent' finite verbs, we need a more specific rule (G2) that overrides the default rule (G1) in the cases stated, i.e. finite verbs selected by German subordinators.

The pre-minimalism constituent based models discussed in section 2 all have difficulties accounting for mixing between SVO and SOV languages because of the opposite setting of the branching parameter. I will show in the next section that this code-mixing pattern can be studied productively in terms of WG.

4. Word Order in Mixed and Monolingual 'Subordinate' Clauses

Code-switching between main and subordinate clauses was chosen as a research area for several reasons. First, it is interesting from a syntactic point of view. If German-English bilinguals want to code-switch subordinate clauses, they need to resolve the problem of English being SVO whereas German finite verbs depending on subordinating conjunctions generally being placed in clause-final position (SOV).[18] How this word order contrast is resolved is relevant to the

underlying question in all grammatical code-switching research, i.e. whether there are syntactic constraints on code-mixing. Second, the code-switched corpus contains a considerable number of switches between main and subordinate clauses (37), not including the 27 switches involving *because* discussed in more detail below. Third, code-switching at clause boundaries has attracted much attention in the research area.

As complementizers are one of the most commonly quoted examples of word classes that are functional categories in constituent-based models of syntax, the government and functional head constraints discussed in section 2, all rule out switching between C and the remainder of the CP. Gumperz (1982) also proposes that subordinate conjunctions must always be in the same code as the conjoined sentence. Sankoff and Poplack (1981: 34), on the other hand, observe that in their Spanish/English corpus subordinate conjunctions tend to remain in the language of the head element on which they depend. Bentahila and Davies' (1983) corpus of Arabic/French yields numerous examples of switches at various types of clause boundary: switches between main clauses and subordinate clauses, switching between complementizers and the clauses they introduce, and examples where the conjunction is in a different language from both clauses.

Although my corpus also contains switches at all the points discussed by Bentahila and Davies (1983), my data largely support Gumperz' (1982) 'constraint', that is, subordinate conjunctions (apart from *because*) tend to be in the language of the subordinate clause that depends on them, and not the head element on which they depend. Examples illustrating switches between main and various types of subordinate clauses in both directions are:

(12) *MEL: ich hab(e) gedacht, **there is going to be a fight.** Jen1.cha, line 987
I have thought

(13) *MEL: **I forgot,** dass wir alle wieder eine neue partie angefangen haben.
that we all again a new game started have
Jen1.cha, line 2541

(14) *TRU: die mutter wird ihr gelernt haben, **how to keep young.**
her mother would her taught have Jen1.cha, line 2016

(15) *DOR: wenn du **short** bist, **you wouldn't talk.**
when you are
*DOR: aber wenn man geld hat, **you talk.**
but when one money has Jen3, line 581-2

(16) *TRU: er schreibt fuenfzehn, **if you leave it in your hand.**
he counts fifteen Jen2.cha, line 932

(17) *LIL: das haengt davon ab, **what 'nasty'** is(t).
that depends on Jen2.cha, line 1062

Note that the null hypothesis is born out in examples (12)–(17) and in the vast majority of monolingual and mixed dependencies[19] in the German–English corpus. The WG rules determining the word order in main and subordinate clauses also hold. These findings are furthermore supported by the quantitative analysis of 1,350 monolingual and 690 mixed dependency relations in a 2,000 word monolingual sample corpus and a 7,000 word code-mixed corpus (see Eppler 2004).

This study particularly focuses on *because* and *weil* clauses. Several researchers (Gardner-Chloros 1991; Salmons 1990; Treffers-Daller 1994; Bolle 1995; Boumans 1998) studying code-mixing between SVO and SOV languages noticed that the clauses depending on switched conjunctions are frequently not SOV but V2. The conjunction in these examples, furthermore, is frequently the causal conjunction *because, parce que* and *omdat*. This led Boumans (1998: 121) to hypothesize that '... it is possible that foreign conjunctions do not trigger verb-final in Dutch and German clauses simply because they are used in functions that require main clause order'. He, however, found it 'hardly feasible to examine this hypothesis in relation to the published examples because these are for the most part presented out of context' (Boumans 1998: 121). I will show that a fully (LIDES[20]) transcribed corpus of German and English data allows us to verify this hypothesis.

Both types of analysis, qualitative structural and quantitative distributional, are considered to be necessary for a comprehensive description of the data, because different structural patterns are used to different degrees and for different purposes. The variation in the data can best be described quantitatively; the qualitative analysis provides an explanation for the structural patterns found. This combination of methodologies furthermore enables us to address Muysken's (2000: 29) statement that '... we do not yet know enough about the relation between frequency distributions of specific grammatical patterns in monolingual speech data and properties of the grammar to handle frequency in bilingual data'. I will compare the *because*- and *weil*-clauses in mixed utterances with monolingual German and English examples and show that we **do** know enough about the syntax and pragmatics of this construction to explain both the frequency distribution of causal conjunctions and the use of verb second (rather than verb final) word order.

4.1 The empirical issues

4.1.1 Asymmetry between conjunctions of reason

The distribution of German and English subordinators/complementizers in the corpus is approximately 60 : 40, which is in accordance with the general distribution of word tokens from the two languages in the data. If, however, we focus on *because* and the translation equivalent from the same word class, the subordinating causal conjunction *weil*, we get a very different picture. The corpus yields twice as many tokens of the English subordinator as it does of *weil* (see Table 1). A typical use of *because*, especially for speaker DOR, is:

(18)	DOR:	es	war	unsere	[...]	Schuld	**because**	man	fühlt	sich
		it	was	our		fault		tone	feels	
		mit	den	eigenen		Leuten	wohler.			
		with	the	own		people	happier.			Ibron.cha, line 221

Because in the above example can be argued to be a single lexical item inserted in otherwise German discourse. This particular usage of the English causal subordinator is not restricted to speaker DOR:

(19)	LIL:	**because**	er	ist	ein	aufbrausender	**Irishman.**
			he	is	a	hot-blooded	
							Jen1.cha, line 389

Because also enters syntactic relations where the word on which it depends is English (*eat*) and its dependent is German (*schmeckt*), as in:

(20)	DOR:	**eat it with**	der Hand-!	**because**	das	schmeckt	ganz	anders.
			the hand		it	tastes	very	differently
								Ibron.cha, line 2214

or vice versa, e.g. *because* has a German head verb (*habe*) but an English complement (*know*):

(21)	MEL:	ich hab's nicht einmal gezählt **because I know I'm going to lose.**
		I have it not even counted
		Jen1.cha, line 881

The German subordinator of reason, *weil*, on the other hand, only enters into monolingual dependency relations:

(22)	DOR:	dann ist sie, weil sie so unglücklich war, dort gestorben.
		then has she, because she so unhappy was, there died
		Ibron.cha, line 1002

So there is not only an asymmetry in the number of tokens each subordinator yields, but also in the language distribution of the immediate syntactic relations which *because* and *weil* enter into, i.e. their main clause head verb and the subordinate dependent verb. The results are summarized in Table 1.

Table 1: Language of head and dependent of *because* and *weil*

	$head_E$ - dep_E	$head_E$ - dep_G	$head_G$ - dep_G	$head_G$ - dep_E	total
Because	86	5	16	6	123
Weil	0	0	59	0	59

The phenomenon of single lexical item subordinate conjunctions in other language contexts is not uncommon in code-mixing literature.[21] As far as directionality of the switch is concerned, the situation in my corpus is in sharp contrast with the findings of Clyne (1973) who studies German/English code-mixing among the Jewish refugee community in Melbourne, Australia. He reports that 'the words transferred from German to English are mainly conjunctions (*denn, ob, und, weil, wie, wo*)' (Clyne 1973: 104). The corpus from the refugee community in London also shows a high propensity for switching

conjunctions, however the vast majority of them are English conjunctions in otherwise German discourse. Lexical transfer of the same word class thus seems to work in the opposite direction in two bilingual communities with a very similar sociolinguistic profile mixing the same language pair.

To rule out the possibility that English *because* is used in place of another German causal conjunction, I will now look at the other possibilities. *Da* is another causal subordinator, thus producing the identical word order effects to *weil*, but normally used in more formal contexts. The whole corpus yields only one example of German *da* used as a subordinating conjunction. This token is embedded in formal discourse and was produced by a speaker who does not use the mixed code as a discourse mode. *Denn* is a causal coordinating conjunction. It was used once by a speaker from the group recordings (not DOR) and three times by a speaker in a more formal setting. *Denn* has increasingly gone out of use in colloquial German (Pasch 1997; Uhmann 1998), however, since it is used by my informants, we need to consider it as a possible translation equivalent of *because*. This possibility is interesting because it involves word order issues: as a coordinating conjunction, *denn* always takes V2 order in the clause following it. The relations between *weil* and *denn* will be discussed further in section 4.2.2 on word order.

4.1.2 Verb second word order after *because* and *weil*

Examples (18)–(20) also demonstrate the structural feature under investigation: German finite verbs occur in main clause word order position in subordinate clauses introduced by *because*. In actual fact not one German finite verb depending on *because* is in clause final position (as in monolingual German subordinate clauses with an overt German subordinator; see example 20).

Furthermore, not all finite dependent verbs follow their subject. Some of them follow fronted indirect objects as in (23), others follow adverbials as in (24):

(23) DOR: **because** dem Computer brauchst' es nicht zeigen.
the computer need you it not show.
Jen2.cha, line 729

(24) LIL: is' wahr -? **because** bei mir hat schon +...[22]
it's true at my place has already
Jen1.cha, line 298

The word order in subordinate clauses after *because* is summarized in Table 2.

Table 2: Word order in subordinate clauses after *because*

	dependent English	dependent German		
	SVX	SVX	XVS	SOV
Because	92	15	6	0

What are supposed to be German dependent verbs occur in second position after *because*, which shows that *because*, at least for my informants, has not taken over the syntactic characteristics of the German subordinating conjunction *weil* which requires its dependent verbs to be clause final.

Let us now take a closer look at this subordinator. Table 1 illustrates that *weil* only has German complements. According to the rules of standard German (rules G2 & G3), finite verbs depending on an overt subordinator should follow all their dependents, i.e. be clause final. This is not borne out in the corpus. Note, however, that 58 per cent of dependent verbs *are* in final position after *weil*, whereas none is in this position after *because*. Table 3 summarizes the position of the dependent finite verb in *weil* clauses from my corpus. In order to see whether verb second after *weil* is a parochial convention of my data or not, I also give the distribution of V2 and Vf from several other corpora of monolingual spoken German[23] for comparison.

Table 3: Verb position after *weil* partly based on Uhmann (1998: 98)

Weil	Vf	V2	Vf	V2
Eppler (2004)	34	25	58%	42%
BYU (Vienna)	62	11	85%	15%
Farrar (1998) BYU	1147	517	69%	31%
Schlobinski (1992)	74	22	70%	23%
Uhmann (1998)	24	19	56%	44%
Dittmar (1997)	99	29	77.3%	22.7%

Table 3 shows that between 15 per cent and 44 per cent of dependent verbs in these corpora are not in final position. So *weil*+V2 word order is not just a peculiarity of the German spoken by my bilingual informants.

We thus have two problems to solve: 1) the asymmetrical distribution of *because* and *weil* in the corpus; and 2) the word order variation in both mixed and monolingual causal clauses introduced by *because* and *weil*. In the next section I will suggest possible solutions to these two problems.

4.2 Possible explanations

4.2.1 For the asymmetry of *because* and *weil*

The frequencies with which *because* and *weil* occur in dependency relations (summarized in Table 1) suggest that for the asymmetry between *because* and *weil* a probabilistic perspective is required.

Fourteen out of the sixteen tokens of *because* in an otherwise German context were produced by one speaker (DOR). This is even more significant if we remember that this speaker is German dominant. The data from this speaker only contain seven tokens of the German subordinator *weil* (and no *denn*). *Because* thus seems to replace *weil* for specific uses in the speech of this speaker. This use of the causal conjunctions is also to be found among the close-knit network of bilinguals who use the mixed code as a discourse mode

(speakers TRU, MEL and LIL); but there is no significant asymmetrical relation between *because* and *weil* in the rest of the corpus.

Reasons for the discrepancy between the British and Australian corpora will have to remain speculative for the moment. I will, however, come back to this point at the end of section 4.2.2. Why German-speaking Jewish refugees in Australia incorporate German conjunctions into their English – and the directionality of lexical transfer being reversed among the same speakers in Britain – could be due to the Australian corpus having been collected approximately 20 years before the London corpus. Michael Clyne collected data from this speech community in the 1970s. My corpus was collected in 1993. An additional two decades of exposure to English of the London-based refugees may be a possible explanation for this discrepancy. Data from American/German dialects that have been in contact with English for up to two centuries support this assumption. See example (25) from Salmons (1990: 472):

(25)	Almost	jedes	moi	is	Suppe	gewen,	**because**	mir	han	kei
	every	time	is	it	soup	be		we	have	no
	Zeit	khat	fer	Supper	recht	essen.				
	time	had	for	soup	properly	to eat				

Treffers-Daller (1994: 192–5) discusses (25) and (26) and suggests analyzing the conjunctions in these two examples as coordinators. For monolingual English Schleppegrell (1991: 323) argues that 'a characterisation of all *because* clauses as subordinate clauses [. . .] is inadequate'. The possibility of a paratactic[24] function for *because* will be discussed in the next section.

Gardner-Chloros's (1991) French/Alsatian data also offer an interesting example of two Alsatian clauses linked by a French causal conjunction.

(26)	Un	noh	isch	de Kleinmann nunter,	**parce que**	ich	hab
	and	now	is	the Kleinmann down	there	I	have
	mi	dort	mue	melde.			
	myself	there	must	check in.			

The German verbs selected by the English and French conjunctions in examples (25) and (26) follow just one dependent, in these cases their subjects. I will discuss the not strictly causal/subordinating use of English *because*, German *weil* and French *parce que* in the next section.

4.2.2 V2 AFTER *BECAUSE* AND *WEIL*

The clearest result of the quantitative analysis presented in Table 2 is that all German finite verbs in clauses after *because* are in second position and none in clause final position.

The Word Grammar rules stated in section 3 account for the empirical data because English subordinators only require finite verbs as their complements (rule E2). German subordinators (rule G3), on the other hand, provide a

specific context that requires dependent verbs to take all their dependents to the left. As *because* is an English subordinator which does not specify that its complement has to be a clause final verb, we get main clause word order (SVO in monolingual English or V2 in mixed utterances).

Supporting evidence for this interpretation comes from the six instances where the finite verb follows a dependent other than its subject (cf. examples 23–24 and 27 below).

(27)	DOR:	**I lost because**	# dreimal	gab	sie	mir	drei	Könige.
			three times	gave	she	me	three	kings
								Jen1.cha, line 817

In the above example the verb is in second position, but the clause is clearly not SVO. The finite verb is preceded by an adverbial but followed by the subject. In other words, the clause displays the V2 order expected in German main clauses.

But how do we know that *because* and the *because*-clause are used in a restrictive subordinating way in examples (23), (24) and (27)? This question needs to be addressed because research conducted by, amongst others, Rutherford (1970), Schleppegrell (1991) and Sweetser (1990), cast doubt on the characterization of *all because*-clauses as subordinate clauses. Especially in spoken discourse, *because* can be used in a variety of non-subordinating and not strictly causal functions.

Several criteria have been proposed to distinguish between restrictive (i.e. subordinating[25]) and non-restrictive *because*-clauses (Rutherford 1970). In sentences containing restrictive *because* clauses yes/no questioning of the whole sentence is possible; pro-ing with *so* or *neither* covers the entire sentence; they can occur inside a factive nominal; and if another *because* clause were added, the two causal clauses would occur in simple conjunction. In semantic terms the main and the subordinate clause form one complex proposition and the *because*-clause provides the cause or reason for the proposition itself. This causal relationship is one of 'real-world' causality (Sweetser 1990: 81). Chafe (1984) asserts that restrictive *because* clauses have a reading that presupposes the truth of the main clause and asserts only the causal relation between the clauses. These clauses tend to have a commaless intonational pattern.

I will now apply these characteristics to some of the causal clauses introduced by *because* in the corpus cited so far. Utterance (27) passes all of Rutherford's (1970) syntactic criteria for restrictive *because*-clauses. The main and *because*-clauses form one complex proposition with a reading in which 'her giving the speaker three kings' is the real world reason for the speaker losing the game of cards. The truth of the sentence-initial clause is presupposed and the causal relation between the two clauses is asserted. These properties of (27) speak for a restrictive analysis. The intonational contour of the utterance, however, displays a short pause after the conjunction.[26] Note furthermore that the causal clause in (27) contains a pre-posed constituent that triggers inversion, i.e. a main clause phenomenon (Green 1976). So there are indicators for both a restrictive/subordinate reading but also syntactic and intonational clues that

point to a non-restrictive/epistemic reading in which the speaker's knowledge causes the conclusion. The latter interpretation suggests non-subordination, which would justify the V2 word-order pattern.

Example (18), repeated here with more context (to facilitate the interpretation) and prosodic information as (28), contains the English conjunction *because* but is otherwise lexified with German words:

(28) DOR: wir waren nie mit richtige Englaender zusammen.
'we never mixed with "real" English people'
DOR: man hätte können # man hat nicht wollen.
'we could have # but we didn't want to'
DOR: *es war unsere [...] Schuld-.*
it was our fault
because *man fühlt sich mit den eigenen Leuten wohler.*
one feels oneself with the own people better
Ibron.cha, line 217-22

This example passes none of Rutherford's (1970) 'tests'. The intonational drop before the conjunction which intonationally separates the two clauses also suggest a non-subordinate analysis for (28). A restrictive reading of the whole construction is awkward if not unacceptable: feeling relaxed in the company of fellow compatriots is not the cause or reason for feeling guilty. The non-restrictive reading in which the *because* clause provides the reason why the speaker said 'it was our own fault' is far more plausible. The *because* clause, furthermore, indicates an interpretative link between clauses that are several utterances apart: the last utterance in (28) provides a 'long-distance' reason for the first utterance in this sequence. Schleppegrell (1991: 333) calls these uses of *because* 'broad-scope thematic links'. They can be only identified when a corpus provides the relevant context for the example. The wider context also identifies the clause preceding the causal clause as presupposed and thematic. The information provided in the causal clause is new and asserted.

The analysis so far suggests that *because* is used in non-restrictive and non-subordinating functions in code-mixed utterances in my corpus. Without repeating them, I will now briefly discuss the other examples in which *because* introduces a clause with predominantly German lexical items (Examples 19–20 and 23–24). Example (19) is a response to a preceding wh-question and thus an independent utterance, the information presented in the reply is not informationally subordinated, it forms the focus of the discourse and provides new information (Schleppegrell 1991: 31). Example (20) has two intonational contours. The intonational rise and the verb first order mark the initial clause as a command or suggestion, i.e. an independent proposition; the following *because* clause then represents an elaboration of that proposition. The content of the causal clause is therefore not presupposed. Example (20) displays all the characteristics of an 'epistemic' (Sweetser 1990) *because*, which indicates 'elaboration and continuation in non-subordinating and non-causal contexts' (Schleppegrell 1991: 323). The *because* clause in example (23) is preceded by a short pause, contains a main clause phenomenon (extraction), and is reflexive

on the previous discourse; finally, the *because* clause in (24) follows a rising intonation of the initial tag, and again explicitly mentions the speaker's knowledge state ('it's true').

We can conclude that those clauses in which *because* has a German (V2) verb as its complement, display more characteristics of 'non-restrictive' (Rutherford 1970) clauses and should therefore be analyzed as paratactic rather than subordinating. The Word Grammar rules formulated in section 3 still account for the data because if *because* is not analyzed as a subordinator, the default rule G1 is not overridden and G2 and G3 do not get activated.

The analysis of the code-mixed data discussed so far indicates that the predominantly German clauses introduced by *because* fulfil functions that are not strictly causal but rather epistemic, broad-scope thematic link, etc. This distinct usage is also reflected in their structural and intonational patterns. We can therefore assume that we are dealing with non-restrictive *because* that is non-subordinate and thus triggers main clause (V2) word order.

However, we also need to consider the monolingual data. The monolingual German data from my corpus are more worrying at first sight. Like *because*, *weil* was traditionally analyzed as a subordinating conjunction with causal meaning which takes a finite verb as its complement. These grammar rules are not absolutely adhered to by my informants and monolingual speakers of German. Only 58 per cent of verbs depending on *weil* in the speech of my informants are in clause final/'late' position. Table 3 shows, furthermore, that in corpora of similar, i.e. southern, varieties of German only 31–85 per cent (with an average of approximately 67 per cent) of the subordinate clauses introduced by *weil* are grammatical according to the rules for monolingual German as stated in section 3.

The recent German literature on *weil* constructions (Günthner 1993, 1996; Pasch 1997; Uhmann 1998), however, suggest an explanation for the monolingual German data and opens up the possibility for an interesting interpretation of the mixed data. There is agreement among the above named researchers that a) there is considerable variation in the use of *weil* + V2 or *weil* + Vf; b) *weil* + V2 is most frequent in southern German dialects; and c) *weil* clauses with verb final placement and *weil* clauses with main clause (V2) word order are found to show systematic but not absolute differences. In a nutshell, the analysis for German *weil* is similar to the analysis proposed for English *because*: there are two types of *weil* clauses, one strictly subordinating one, and several non-restrictive paratactic uses. The factor that best seems to account for the data is the information structure of the construction. If pragmatics and syntax, which in German is a much clearer indicator than in English, fail to provide clear criteria as to which type of *weil*-construction we are dealing with, intonation can once again help to disambiguate. Example (29) from my corpus illustrates epistemic *weil* + V2:

(29) LIL: sie hat sich gedacht, die [/] die muss doch Wien kennenlernen,
'She thought she needs to get to know Vienna'
weil die eltern sind beide aus Wien.
because parents are both from Vienna
Jen3.cha, line 107-8

Note that in (29) *weil* could be replaced by the German coordinating conjunction *denn*. Pasch (1997) and Uhmann (1998) agree that the non-restrictive *weil* seems to take the position of Standard German *denn* in the system of conjunctions of reason in colloquial German.

In the analysis so far it has been established that there are 'restrictive' and 'non-restrictive' *because* clauses in English and 'restrictive' and 'non-restrictive' *weil* clauses in German. A cross-linguistic comparison of these clause types revealed that they share many of their discourse-pragmatic, syntactic and intonational characteristics. My informants use both clause types from both languages in monolingual contexts. In addition to this, they employ *because* in code-mixed contexts. They treat English *because* as the translation equivalent of the non-restrictive *weil*+V2 or *denn*. Their linguistic competence tells them that these constructions are equivalent in syntax and pragmatic content.

This was demonstrated for the quoted examples and also holds true for the *because* followed by weil+V2 examples not reproduced in this chapter. Furthermore, if we apply this analysis to the quantitative asymmetry found in the corpus between the two conjunctions *because* and *weil* and add the 21 tokens of *because*+V2 to the *weil* tokens, this asymmetry shrinks to a figure (80 *weil* : 120 *because*) which is in line with the general language distribution in the corpus. In addition to the syntactic and pragmatic reasons for using this 'congruence approach' (Sebba 1998: 1) to switching at clause boundaries, my informants may also be dialectally pre-disposed to the *weil*+V2 construction because all of them are L_1 speakers of a southern German variety.

I will now briefly return to the discrepancy between the Australian (Clyne 1973) and London corpora mentioned in sections 4.1.1 and 4.2.1. The question was why German speaking Jewish refugees in Australia incorporate German conjunctions into their English, and the directionality of lexical transfer is reversed among the same speakers in Britain. I hypothesized that duration of language contact may have something to do with it. At the time of data collection, German speaking refugees in Australia had been mixing German and English for approximately 30 years. In London, on the other hand, these two languages had been in contact for more than half a century when I collected my data. Another situation where we can witness long-term contact between the two languages under investigation are German-American dialects. Note, furthermore, that example (25) from these data (Salmons 1990) also has main clause word order after *because*.

The development in Pennsylvania German (Louden 2003) is particularly interesting in this respect. Louden (2003) illustrates the causal conjunction paradigm in Pennsylvania German (PG) data from the 19th century onwards. In the second half of this century he found the standard German distribution of *weil* + verb final and *dann* (< Germ. *denn)* + V2. In data from the beginning of the 20th century PG still has verbs depending on *weil* in final position; *dann*, however, has been replaced by *fer* (< Engl. *for*) + V2. In modern sectarian PG *weil* is backed up with *(d)ass*, a historical merger of *dass* with comparative *als*, and *for* (originally *dann* < Germ. *denn*) has been replaced with *because* + V2.

This development is interesting for several reasons: PG, in the late 19th, early 20th century went through a phase that mirrors present-day English in the distribution between *because* and *for*. In modern Pennsylvania German, *weil* does not seem to be able to function as subordinator in its own right any longer and it has to be backed up by another complementizer to trigger verb final placement. This supports rule G2 (section 3) which implicitly proposes a subordinate feature on lexical complementizers. Modern PG seems to have lost this feature and therefore needs to be 'backed up' by another subordinator to trigger verb final word order.

Dann in modern PG, on the other hand, after having gone through the stage of *fer* (<Engl. *for*), is eventually replaced by *because*, as in my data. This development not only backs up the speculation voiced in section 4.1.2, i.e. that the discrepancy between my and Clyne's (1973) German-English corpora might be due to prolonged language contact, but also the qualitative analysis presented in section 4.2.2.

The WG stipulation of a subordinate feature on German complementizers (Rule G3) is furthermore supported by data from another language contact situation with word order contrasts in subordinate clauses: French Dutch contact in Brussels. The most frequently borrowed subordinator in Brussels Dutch is *tandis que*. Treffers-Daller (1994: 191) observes that the Dutch equivalent of *tandis que*, *terwijl*, is rarely used in her corpus. In those cases that do occur, the Dutch conjunction is followed by the Dutch complementizer *dat*. Like *weil* in Pennsylvania German, Brussels Dutch *terwijl* may also have lost the subordinate feature and require an 'obvious' complementizer to trigger verb final placement.

5. Summary and Conclusion

In section 2 of this chapter I illustrated why syntactic constraints on intra-sentential code-mixing formulated within Phrase Structure Grammar frameworks are empirically not borne out. They are too restrictive because the domain of government, i.e. maximal projections or constituents, was too large, and because of the problematic distinction between lexical and functional categories.

In section 3 I discussed the advantages of WG over other linguistic theories for code-mixing research. They are seen to be:

- Word Grammar requires less analysis than constituency-based models because the only units that need to be processed are individual words. Larger units are built by dependency relations between two words which can be looked at individually.
- As syntactic structure consists of dependencies between pairs of single words, constraints on code-mixing are less prone to over-generalization than constraints involving notions of government and constituency.
- Word Grammar allows a single, completely surface analysis (with extra dependencies drawn below the sentence-words). Code-mixing seems to be a surface-structure phenomenon, so this property of WG fits the data.

- Knowledge of language is assumed to be a particular case of more general types of knowledge. Word Grammar accommodates sophisticated socio-linguistic information about speakers and speech communities. This is important for language contact phenomena that are influenced by social and psychological factors.
- In contrast with most other syntactic theories, Word Grammar recognizes utterances.
- WG is a competence model which can handle inherent variability.

I do not claim that the present work illuminates theories of language structure but it confronts a linguistic theory, Word Grammar, with statistical data, and shows that this theory of language structure can be successfully and illuminatingly used for the analysis of monolingual and code-mixed constructions. The WG formulation of the null hypothesis is born out with just a handful of exceptions, and the WG rules determining word order in monolingual German or English and code-mixed clauses also hold.

The investigation of word order in subordinate clauses, furthermore, shows that the null hypotheses seems to be correct even in cases where we would expect restrictions on code-switching due to surface word order differences between the two grammars involved in mixing. The quantitative analysis of monolingual and code-mixed *because* and *weil* clauses revealed that a) the core group of informants favour the English causal conjunction *because* over German *weil* or *denn*; the use of *weil* and *denn* are restricted to monolingual German contexts, and *because* is also used to introduce mixed utterances; b) the word order in *weil* clauses varies between verb final, as required in subordinate clauses, and verb second, the main clause order; the coordinating conjunction *denn* only occurs once and with main clause order, as expected; mixed clauses introduced by *because* invariably have verb second structure. Independent research on the syntactic, intonational, semantic and pragmatic properties of monolingual *because* and *weil* clauses has shown that these properties cluster to form two main types of causal clauses: restrictive and non-restrictive (Rutherford 1970). The qualitative analysis of the monolingual causal clauses in the corpus revealed that they also fall into these two types *and* that the mixed utterances introduced by *because* predominantly have the grammatical properties of non-restrictive clauses. Thus Boumans' (1998: 121) hypothesis that 'foreign conjunctions do not trigger verb-final in German clauses simply because they are used in functions that require main clause order' could be verified. The quantitative analysis of *because* and *weil* clauses has furthermore demonstrated how frequency distributions of a specific grammatical pattern in monolingual speech data can be combined with our knowledge about syntactic and pragmatic properties of grammars to handle frequency in bilingual data (Muysken 2000).

The WG analysis of German (and Dutch) lexical subordinators having a 'subordinate' feature which triggers verb final placement was furthermore supported by data from two other language contact situations (Pennsylvania German and Brussels Dutch) in which certain subordinators seem to have lost this feature and therefore to require 'backing up' from overt complementizers.

References

Belazi, H. M., Rubin, E. J. and Toribio, A. J. (1994), 'Code switching and X-bar theory: The functional head constraint'. *Linguistic Inquiry*, 25, 221–37.

Bentahila, A. and Davies, E. E. (1983), 'The syntax of Arabic – French code – switching'. *Lingua*, 59, 301–30.

Bolle, J. (1995), 'Mengelmoes: Saranan and Dutch language contact', in *Papers from the Summer School Code-switching and Language Contact*. Ljouwert/Leeuwarden: Fryske Akademie, pp. 290–4.

Boumans, L. (1998), *The Syntax of Codeswitching: Analysing Moroccan Arabic/Dutch Conversations*. Tilburg: Tilburg University Press.

Chafe, W. L. (1984), 'How people use adverbial clauses'. *Berkeley Linguistics Society*, 10, 437–49.

Chomsky, N. (1981), *Lectures on Government and Binding*. Dordrecht: Foris.

Clyne, M. G. (1973), 'Thirty years later: Some observations on "Refugee German" in Melbourne', in H. Scholler and J. Reidy (eds) *Lexicography and Dialect Geography, Festgabe for Hans Kurath*. Wiesbaden: Steiner, pp. 96–106.

– (1987), 'Constraints on code-switching: how universal are they?' *Linguistics*, 25, 739–64.

DiSciullo, A-M., Muysken P. and Singh, R. (1986), 'Government and Code-Mixing'. *Journal of Linguistics*, 22, 1–24.

Eppler, E. (1999), 'Word order in German-English mixed discourse'. *UCL Working Papers in Linguistics*, 11, 285–308.

– (2004), '... Because dem Computer brauchst' es ja nicht zeigen': Because + German main clause word order'. *International Journal of Bilingualism*, 8, 127–43.

– German/English LIDES database <*http://talkbank.org/data/LIDES/Eppler.zip*>.

Gardner-Chloros, P. (1991), *Language Selection and Switching in Strasbourg*. Oxford: Clarendon Press.

Green, G. M. (1976), 'Main clause phenomena in subordinate clauses'. *Language*, 52, 382–97.

Grosjean, F. (1995), 'A psycholinguistic approach to codeswitching', in P. Muysken and L. Milroy (eds), *One Speaker, Two Languages*. Cambridge: Cambridge University Press, pp. 259–75.

Gumperz, J. J. (1982), *Discourse Strategies*. Cambridge: Cambridge University Press.

Günther, S. (1993), '...Weil-man kann es ja wissenschaftlich untersuchen'- Diskurspragmatische Aspekte der Wortstellung in weil-Sätzen'. *Linguistische Berichte*, 143, 37–59.

– (1996), 'From subordination to coordination?'. *Pragmatics*, 6, 323–56.

Hudson, R. A. (1980), *Sociolinguistics*. Cambridge: Cambridge University Press.

– (1997), 'Inherent variability and linguistic theory'. *Cognitive Linguistics*, 8, 73–108.

– (1990), *English Word Grammar*. Oxford: Blackwell.

– (2000), 'Grammar without functional categories', in R. Borsley (ed.), *The Nature and Function of Syntactic Categories*. New York: Academic Press, pp. 7–35.

Joshi, A. K. (1985), 'Processing of sentences with intrasentential code-switching', in L. Dowty, L. Kartunnen and A. M. Zwicky (eds), *Natural Language Parsing*. Cambridge: Cambridge University Press, pp. 190–205.

Lehmann, Ch. (1988), 'Towards a typology of clause linkage' in J. Haiman and S. Thompson (eds), *Clause combining in grammar and discourse*. Amsterdam/Philadelphia: John Benjamins, pp. 181–226.

Louden, M. L. (2003), 'Subordinate clause structure in Pennsylvania German'. *FGLS/SGL Joint Meeting*. London, 2003.

MacSwan, J. (1999), *A Minimalist Approach to Intrasentential Codeswitching*. New York and London: Garland.
Mahootian, S. and Santorini, B. (1996), 'Code-switching and the complement/adjunct distinction'. *Linguistic Inquiry*, 27, 3, 464–79.
Muysken, P. (1989), 'A unified theory of local coherence in language contact', in P. Nelde (ed.), *Language Contact and Conflict*. Brussels: Centre for the Study of Multilingualism, pp. 123–9.
– (2000), *Bilingual Speech. A Typology of Code-Mixing*. Cambridge: Cambridge University Press.
Myers-Scotton, C. (1993), *Duelling Languages: Grammatical Structure in Code-Switching*. Oxford: Oxford University Press.
Myers-Scotton, C. and Jake, J. L. (1995), 'Matching lemmas in a bilingual language competence and production model: Evidence from intrasentential code switching'. *Linguistics*, 33, 981–1024.
Pasch, R. (1997), '*Weil* mit Hauptsatz-Kuckucksei im denn-Nest?'. *Deutsche Sprache*, 25, 252–71.
Poplack, S. (1980), 'Sometime I'll start a sentence in Spanish y termino en Espanol'. *Linguistics*, 18, 581–618.
Poplack and Meechan (1995), 'Orphan categories in bilingual discourse: A comparative study of adjectivization strategies in Wolof/French and Fongbe/French'. *Language Variation and Change*, 7, 2, 169–94.
Romaine, S. (1989), *Bilingualism*. Malden, Mass.: Blackwell.
Rutherford, W. E. (1970), 'Some observations concerning subordinate clauses in English'. *Language*, 46, 97–115.
Salmons, J. (1990), 'Bilingual discourse marking: Code switching, borrowing, and convergence in some German-American dialects'. *Linguistics*, 28, 453–80.
Sankoff, D. and Poplack, S. (1981), 'A Formal Grammar for Code-Switching'. *Papers in Linguistics*, 14, 3–46.
Schleppegrell, M. J. (1991), 'Paratactic *because*'. *Journal of Pragmatics*, 16, 323–37.
Schlobinski, P. (1992), 'Nexus druch *weil*', in P. Schlobinsky (ed.) *Funktionale Grammatik und Sprachbeschreibung*. Opladen: Westdeutscher Verlag. 315–44.
Scotton, C. M. (1990), 'Code-switching and borrowing: Interpersonal and macro-level meaning', in R. Jacobson (ed.), *Codeswitching as a Wordwide Phenomenon*, New York: Peter Lang, pp. 85–110.
Sebba, M. (1998), 'A congruence approach to the syntax of codeswitching'. *International Journal of Bilingualism*, 2, 1–19.
Sweetser, E. (1990), *From Etymology to Pragmatics. Metaphorical and Cultural Aspects of Semantic Structure*. Cambridge: Cambridge University Press.
Thorne, J. P. (1986), 'Because', in D. Kastovsky and A. Szwedek (eds) *Linguistics across Historical and Geographical Boundaries* (Vol. 12). Berlin: Mouton. 1063–6.
Treffers-Daller, J. (1994), *Mixing Two Languages: French-Dutch Contact in a Comparative Perspective*. Berlin: de Gruyter.
Uhmann, S. (1998), 'Verbstellungsvariationen in weil-Sätzen'. *Zeitschrift für Sprachwissenschaft*, 17, 29–139.
Zwicky, Arnold (1992), 'Some choices in the theory of morphology', in Robert Levine (ed.) *Formal Grammar: Theory and Implementation*. Oxford: Oxford University Press, pp. 327–71.

Notes

1 The corpus was collected in 1993 from German-speaking Jewish refugees residing in London. All transcripts are available on <http://talkbank.ord/data/LIDES/Eppler.zip>.
2 Categorial equivalence is 'when the switched element has the same status in the two languages, is morphologically encapsulated, shielded off by a functional element from the matrix language, or could belong to either language' (Muysken 2000: 31).
3 Myers-Scotton and Jake (1995: 985) define system morphemes as morphemes that do not participate in the thematic structure of a sentence, i.e. they are specified as [- theta-role assigner/receiver]. A second feature characteristic of 'most' system morphemes is the feature [+ Qantification]. A morpheme has a plus setting for quantification within the Matrix Language Frame model, if it restricts possible referents of a lexical category. Myers-Scotton and Jake (1995: 985) give tense and aspect as examples for [+ Q]. Tense and aspect restrict the possible reference of predicates (i.e. verbs and adjectives). Prototypical system morphemes are inflections and most function words.
4 The WG approach of incorporating different syntactic properties of WORDS in isa-hierarchies seems more economical and convincing.
5 The Free Morpheme Constraint (Sankoff and Poplack 1981: 5) prohibits switching between a bound morpheme (pre- or suffix) and a lexical form unless the latter has been phonologically integrated into the language of the bound morpheme.
6 Note the similarity of this corollary with the WG null hypothesis this study is based on.
7 Constituency analysis is applied only to coordinate structures.
8 This system implies that code-mixing ought to be less frequent among typologically quite different language pairs.
9 According to the theory of markedness (Scotton 1990: 90), speakers know that for a particular conventionalized exchange in their community, a certain code choice will be the unmarked realization of an expected rights and obligations set between participants. They also know that other possible choices are more or less marked because they are indexical of other than the expected rights and obligations set.
10 Smooth code-switches are unmarked by false starts, hesitations, lengthy pauses, etc.; flagged switches are accompanied by discourse markers and other editing phenomena (Poplack 1980).
11 A database is particularly important for studies of codes that do not have 'native' speakers who can provide fairly reliable grammaticality judgements. A corpus is also an essential test for the constraints on and models of code-mixing.
12 An alternative analysis of this example would be that it is ambiguous, i.e. it conforms to two different models. The stem conforms to the English phonological model and the suffix conforms to the German plural suffix; i.e. it is a morphologically integrated borrowed stem.
13 The figures for individual cases need not be the same; cases of lexical diffusion would seem to suggest the contrary (Hudson 1980: 168ff). And presumably the entrenchment value for the general rule in such cases could be different from all the individual rules.
14 Default inheritance rules apply to the few English constructions in which the complement comes before the head.
15 These rules are not intended to cover scrambling, double infinitive constructions and other well-known word order intricacies of German.

16 The term 'late' was chosen instead of 'final' because finite dependent auxiliaries in double infinitive constructions can be followed by their non-finite dependents; cf. endnote 15.

17 Support for this analysis comes from the fact that German subordinate clauses lacking a subordinator/complementizer are V2 (or verb initial). Cf.:

Sie	*sagte,*	*sie*	*kennen*	*Doris*	vs.	*Sie*	*sagte,*	*daß*	*sie*	*Doris kennen*
She	said	they	know	Doris		She	said	that	they	Doris know

According to G3, it is only subordinators/complementizers that select 'late' finite verbs. So if a verb depends directly on another verb (*kennen* directly depending on *sagte* and not *daß*) the default rule need not be overridden.

18 Exceptions to this rule are extraposition and double-infinitive constructions.

19 The null hypothesis is violated in five tokens of two construction types: word-order violations of objects and negatives (see Eppler 2004).

20 The data this study is based on are transcribed in the LIDES (Language Interaction Data Exchange) system. More information on the transcription system can be found on <www.ling.lancs.ac.uk/staff/mark/lipps/ >.

21 See for example Clyne (1987), Gardner-Chloros (1984), Salmons (1990), Treffers-Daller (1994).

22 Example (24) is an incomplete subordinate clauses. This does not effect the analysis because the word order position of the relevant finite dependent verb is clear.

23 Since all my informants are from Vienna, I used only examples from the ten Viennese informants for the Brigham Young Corpus (BYU) corpus. Farrar (1998) counted all occurrences of *weil* in the speakers of southern German dialects from the BYU corpus. Schlobinski's (1992) data are standard Bavarian; and the Uhmann (1998) corpus is 'alemannisch-bairisch'.

24 Lehmann (1988) suggests that for clauses that are linked in a relationship of sociation rather than dependency, 'paratixis' is a more appropriate term than 'coordination'.

25 Two clauses (X and Y) have been defined as being in a subordination relationship 'if X and Y form an endocentric construction with Y as the head' (Lehmann 1988: 182).

26 Note that in the English literature, Rutherford (1970) and Thorne (1986), the comma intonation is assumed to precede the conjunction. Schleppegrell (1991: 333) mentions the possibility of *because* followed by a pause.

7 Word Grammar Surface Structures and HPSG Order Domains*

TAKAFUMI MAEKAWA

Abstract

In this chapter, we look at three different approaches to the asymmetries between main and embedded clauses with respect to the elements in the left periphery of a clause: the dependency-based approach within Word Grammar (Hudson 2003), the Constructional Head-driven Phrase Structure Grammar (HPSG) approach along the lines of Ginzburg and Sag (2000), and the Linearization HPSG analysis by Chung and Kim (2003). We argue that the approaches within WG and the Constructional HPSG have some problems in dealing with the relevant facts, but that Linearization HPSG provides a straightforward account of them. This conclusion suggests that linear order should be independent to a considerable extent from combinatorial structure, such as dependency or phrase structure.

1. Introduction

There are two ways to represent the relationship between individual wor DEPENDENCY STRUCTURE and PHRASE STRUCTURE. The former is a p representation of word-word relationships while the latter includes additio information that words are combined to form constituents. If all work can done just by means of the relationship between individual words, phr structure is redundant and hence dependency structure is preferable to it would therefore be worth considering whether all work can really be done v just dependencies. We will look from this perspective at certain linear or asymmetries between main clauses and subordinate clauses. One example such asymmetries can be seen in the contrast of (1) and (2). The former sh that a topic can precede a fronted *wh*-element in a main clause:

(1) a. Who had ice-cream for supper?
b. For supper who had ice-cream?

(2) illustrates, however, that this is not possible in an embedded clause:

(2) a. Who had ice-cream for supper is unclear.