Prince, A. & Smolensky, P. (1993). Optimality Theory: constraint interaction in Generative Grammar. Technical Report, Rutgers University and University of Colorado, Boulder. (Also available as ROA-537, Rutgers Optimality Archive, http://roa.rutgers.edu/.)
Zonnefeld, W. (1998). Lutgart, Willem and Geoffrey. A study of 13th-century Dutch meter. Ms., University of Utrecht.

Author's address: *Department of English Language and Literature, Edge Hill College of Higher Education, Ormskirk, Lancashire L39 4QP, U.K.*
E-mail: honeybop@edgehill.ac.uk

(Received 27 January 2003)

J. Linguistics **39** (2003). DOI: 10.1017/S0022226703272297

Pieter Muysken, *Bilingual speech: a typology of code-mixing*. Cambridge: Cambridge University Press, 2000. Pp. xvi + 306.

Reviewed by SHANA POPLACK, University of Ottawa & JAMES A. WALKER, York University

In the contentious world of code-mixing (CM)[1] research, where most contributions are little more than attempts to discredit earlier work, and each successive model proclaims universal applicability to all existing and future bilingual data, Pieter Muysken's has always been the voice of reason. In contrast to the prevailing emphasis on the uniqueness of code-mixing theories, Muysken's efforts have been directed to understanding how they resemble each other, and where (and why) their predictions overlap. *Bilingual speech* is the culmination of over twenty years of such efforts to make sense of the diverse and often contradictory CM literature, viewed through the lens of a tripartite division of CM that Muysken views as his 'main contribution' (32): INSERTION of material from one language into structure from the other language, ALTERNATION between the structures of the two languages and CONGRUENT LEXICALIZATION (CL) of material from different lexicons into a shared grammatical structure.[2] The goal is modest (perhaps necessarily so, given the state of the field): to 'tie together a set of intermediary results rather than giving a conclusive account' (2).

In the first chapter, 'The study of code-mixing', Muysken provides an overview of research on language mixture. He argues that the various

[1] Muysken uses the term 'code-mixing' to refer to 'all cases where lexical items and grammatical features from two languages appear in one sentence' (1), restricting the term 'code-switching' to a subset of CM. For the purposes of this review, we follow Muysken's terminology.

[2] One of the dominant traditions in CM research distinguishes insertion from alternation, in contrast to unitary theories that attempt to provide a single analysis for all CM. The three-way division (and the phenomena to be included under each) is original to Muysken.

grammatical constraints on CM that have been proposed (e.g. Poplack's (1980) Free Morpheme and Equivalence Constraints; di Sciullo, Muysken & Singh's (1986) Government Constraint; Myers-Scotton's (1993) Matrix Language Framework) can be characterized in terms of each of the three CM processes. Noting that the field has moved from constraints specific to particular language pairs to more universal principles, he argues further that all constraints can be reduced to a number of 'primitives' involving issues of equivalence (categorial or syntactic), clausal peripherality and the role of function words. Muysken makes the bold claim that CM is 'impossible in principle' (30) and views the three CM strategies as 'escape hatches' within a unified theory of bilingual speech.

The second chapter, 'Differences and similarities between languages', provides theoretical perspectives on language differences and on the division of labor between grammar and the lexicon. If all such differences are ultimately lexical, in many cases reflecting the requirements of function words, potential violations of CM processes can be attributed to the (in)compatibility of function words in the two languages. Reviewing the mismatches between lexical and grammatical structures as dimensions in typological classification, Muysken arrives at his working hypothesis that differences between languages result from differences in the interaction of different autonomous modules: specifically, whether information is encoded lexically or grammatically (51).

In the next three chapters, Muysken discusses each of the three CM strategies individually, providing a wealth of examples culled from a broad range of studies. Chapter 3, 'Insertion', examines the grammatical dimensions of INSERTIONAL CM, uniting lexical borrowing, nonce borrowing (Sankoff, Poplack & Vanniarajan 1990) and constituent insertion (Naït-M'barek & Sankoff 1988). Insertions, which are morphologically integrated lexical (rather than functional) elements, form a single syntactic constituent, usually an object or complement (rather than an adjunct) which exhibits a nested structure ($L_A [_A L_B]$). Insertion implies the existence of a base or matrix language (ML) and considerations of syntactic dependency, but Muysken concludes that problems in determining the ML are empirical rather than theoretical (68). He invokes the notion of government (di Sciullo et al. 1986) to account for observed selectional restrictions. Although insertional CM bears obvious similarities to lexical borrowing, Muysken insists that the former is 'supralexical', while the latter is 'sublexical' and 'listed' (i.e. 'part of a memorized list which has gained acceptance within a particular speech community' (71)).

Chapter 4, 'Alternation', discusses the properties of ALTERNATIONAL CM. Unlike insertion, alternation involves the switch of longer, more complex elements, typically multiword constituents in a non-nested sequence ($L_A \ldots L_B$). Alternation is a process characterized by the absence of selectional restrictions (other than equivalence) in which clausally peripheral

elements such as adverbials and discourse particles often figure. Reviewing CM models formulated in terms of phrase structure, Muysken concludes that linear equivalence is better conceptualized as a subset of categorial equivalence and that equivalence constraints should be oriented to surface structures rather than deep structures, in contrast with models that adhere to Chomskyan syntactic theories.

The consequences of CM for grammatical convergence and linguistic variation are outlined in chapter 5, 'Congruent lexicalization'. Unlike insertion and alternation, which impose strict grammatical requirements on CM, 'anything goes' in CL (128). Constituents may be single or multiple (or not even constituents at all) and may belong to any category (lexical or functional). CL is bidirectional, characterized by back-and-forth switching and the existence of 'diamorphs' (words that are homophonous in the two languages).

Muysken examines in more detail the importance of function words in CM (the 'functional element effect') in chapter 6, 'Function words', and argues that their restricted participation in CM is a result of the non-equivalence of function words across languages rather than considerations of language production (contra Myers-Scotton 1993). Reviewing the different definitions of function words, Muysken concludes that the functional-lexical distinction is gradient and proposes a distinction couched in terms of cross-linguistic equivalence. According to Muysken, nouns and verbs are universal, but the featural complexes associated with each are language specific. For example, the insertion of an L_A verb triggers the L_A tense system.

Chapter 7, 'Bilingual verbs', considers the phenomenon of verbal compounds combining elements from two languages, manifested in different ways: unmarked verbs, verbs marked with native affixes and verbs adapted (morphologically or phonologically) before being morphosyntactically integrated. Muysken argues (215) that we cannot adopt a unitary analysis of bilingual verbs but rather must classify them into three types, corresponding to each of his CM strategies: nominalized verbs (insertion), adjoined verbs (alternation) or combinations of auxiliary and infinitive (CL).

In chapter 8, 'Variation in mixing patterns', Muysken relates the different CM strategies to psycholinguistic and social factors: language dominance, duration of contact, bilingual proficiency, speaker type, age-group or generation and language attitudes. He associates insertion with shorter-duration contact situations and more isolated groups (such as recent immigrant communities); alternation with communities with strong norms, competition between language groups and typological distance between languages; and CL with looser norms, balanced bilingualism and structurally parallel languages. Nonetheless, after attempting to characterize a number of bilingual communities according to CM strategies, Muysken correctly concludes that differences between them are not absolute, since various social factors can interact to yield mixed CM strategies within each community.

A final chapter, 'Code-mixing, bilingual speech and language change', relates CM strategies to bilingual production and to language change. Muysken argues that bilingual speech data constitute evidence for a 'simultaneous access' model of production, with different modules from each language activated. He cites Thomason & Kaufman (1988) in considering the relevance of CM strategies to processes of language contact, such as relexification, convergence, pidginization, lexical borrowing, second language acquisition and substratum effects. These processes, he suggests (268), can be viewed as the gradual importation of more and more structure from one language into another. Muysken concludes with a discussion of future avenues of research, such as the difference between content words and function words, the role of adjunction as a 'fallback' strategy ('potentially available … at moments when the grammar fails' (277)) and considerations of language separation and economy in interference.

To what extent will the model proposed in this ambitious volume succeed in providing a framework for the study of CM, or even a 'taxonomic phase' (2) in the study of CM data? This remains to be seen. While Muysken provides detailed theoretical justification and exemplification for his proposal to distinguish three different CM strategies, they have yet to be applied SYSTEMATICALLY to corpora of actual bilingual production. Insertion and alternation will be familiar to most readers, whether or not they subscribe to the distinction between borrowing and code-switching. The concept of 'congruent lexicalization' is more elusive, though likened by Muysken to monolingual (e.g. stylistic) variation. Muysken presents an eloquent plea to conflate CM and linguistic variation, 'to deal with variation and its quantitative coordinates in terms of variation in lexical insertion' (126). How can such an analysis be implemented? No heuristic is provided. Indeed, readers searching for empirical criteria to distinguish among the three CM types will be disappointed, since the same criteria are often offered for more than one type. For example, morphological integration is diagnostic of both insertion (64) and CL (134); CL may occur in borrowing, the main type of insertion (123); and phonological integration is variable in borrowing as well as in other types of CM (70). Muysken is not unaware of this dilemma, observing that 'it is impossible to *prove*, for every case, that it is alternation, insertion or CL. The best we can do is study patterning at the level of the whole corpus' (231). Identification of patterns requires quantitative analysis, raising the issue of method.

Muysken's methods are as eclectic as his ideas. He states at the outset that he will combine structural analysis with 'quantitative analysis as in the work of Labov and Sankoff' (2). Correctly rejecting the rule-and-exception paradigm so characteristic of CM theories, he endorses instead 'probabilistic statements, linked to different language pairs and contact settings' (28). A good deal of the discussion does in fact refer to percentages, many reproduced or adapted from the quantitative work of others. His examination

(chapter 8) of patterns of co-occurrence of CM types in different bilingual communities (or more accurately, data sets) relies particularly heavily on quantitative trends. He rightly concludes (247) that the variation in mixing patterns is not explicable by a single factor, a situation which lends itself perfectly to multivariate analysis. Muysken's assertion that a probabilistic model is 'only possible in the abstract at present' (249) is puzzling, given the large body of empirical work that has successfully implemented this model, e.g. in predicting syntactic sites likely to host a switch, in a preference for different CM types in different bilingual communities, and in distinguishing different language contact phenomena. By the volume's end, he has dismissed this approach: 'we have reached the limits of what can be learned about [CM] using the Labovian techniques introduced in the 1960s' (250). Yet most of the limitations he enumerates, such as lack of diachronic perspective and the primacy of syntax (250), result from the prevailing research climate rather than the Labovian/quantitative paradigm, which is amply equipped to handle them. (Indeed, one could argue that the reason CL remains the weakest link in Muysken's tripartite division is precisely because its nature (and even its existence) have not been subjected to the rigors of the variationist method.) An exclusive 'reliance on corpora of spontaneous bilingual speech' (250) is endemic to variationism, a fact which Muysken laments – curiously, since the whole of chapter 8 (and much of the rest of the volume) would have been impossible without such corpora. His call for greater reliance on experimental data (249) contradicts the problems he cites (28f.) as inherent to them, and places unwarranted faith in the ability of elicitation and experimentation to resolve the outstanding problems in his analysis.

In short, while there is much to quibble with here, this should not obscure the remarkable achievements of this original and exciting book. Its scope, cogent argumentation and vast range of examples are all testaments to the breadth of Muysken's interests. He has undertaken the daunting task of confronting the formidable literature on CM and has attempted to ground its findings in research from diverse fields. Perhaps the greatest contribution of this volume is its success in conveying 'the excitement of working in a field that is moving quite rapidly and is located at the crossroads of structural analysis, sociolinguistics and psycholinguistics' (278).

REFERENCES

di Sciullo, A.-M., Muysken, P. & Singh, R. (1986). Government and code-mixing. *Journal of Linguistics* **22**. 1–24.

Myers-Scotton, C. (1993). *Duelling languages: grammatical structure in codeswitching*. Oxford: Oxford University Press.

Naït-M'barek, M. & Sankoff, D. (1988). Le discours mixte arabe/français: emprunts ou alternances de langue? *Canadian Journal of Linguistics* **33**. 143–154.

Poplack, S. (1980). 'Sometimes I'll start a sentence in Spanish *y termino en español*': toward a typology of code-switching. *Linguistics* **18**. 581–618.

Sankoff, D., Poplack, S. & Vanniarajan, S. (1990). The case of the nonce loan in Tamil. *Language Variation and Change* **2**. 71–101.
Thomason, S. & Kaufman, T. (1988). *Language contact, creolization, and genetic linguistics*. Berkeley & Los Angeles: University of California Press.

Authors' addresses: (*Poplack*) *Department of Linguistics, University of Ottawa, 70 Laurier Avenue E., #401, Ottawa, ON, Canada K1N 6N5. E-mail: spoplack@aix1.uottawa.ca*
(*Walker*) *Department of Languages, Literatures and Linguistics, York University, 4700 Keele St., Toronto, ON, Canada M3J 1P3. E-mail: jamesw@yorku.ca*

(Received 4 February 2003)

J. Linguistics **39** (2003). DOI: 10.1017/S0022226703282293

Henk van Riemsdijk (ed.), *Clitics in the languages of Europe* (Empirical Approaches to Language Typology). Berlin: Mouton de Gruyter, 1999. Pp. xxii + 1026.

Reviewed by CATHERINE RUDIN, Wayne State College

The first striking fact about this volume is its sheer bulk: at well over 1,000 pages, surely it must contain everything you ever wanted to know about clitics? And indeed, it does provide considerable depth and breadth of information, along with a sense of the ongoing theoretical controversies surrounding clitics. The book caps five years of collaboration by one of the working groups of the 'Typology of Languages in Europe' (EUROTYP) project. As such, it brings together the work of numerous experts in an unusually coherent format, with interlinking among chapters and a rare degree of give and take among the authors, providing an excellent overview of the state of clitic studies in the late 1990s.

After the editor's introduction, the volume is organized into two unequal parts, plus a 118-page appendix . Part I, 'Area studies', contains three overview articles totaling 110 pages. Part II, 'Theory', is 742 pages long. It consists of two 'Feature articles', each with peer comments by several scholars and a reply by the author(s), and five 'Topics' sections: clusters of one to five articles on related issues. The volume also contains lists of abbreviations and of contributors' addresses, two prefaces, and three indexes (languages, names and subjects).

Van Riemsdijk's introductory article, 'Clitics: a state-of-the-art report', effectively sets the background for the volume, summarizing major viewpoints on the properties and analysis of clitics from Kayne (1975) and Zwicky (1977) through much more recent work.

The 'area' (actually family/subfamily) overview articles in Part I are by Anna Cardinaletti on Germanic and Romance, Mila Dimitrova-Vulchanova