

Morphosyntactic Convergence Versus Pragmatic Divergence: A Cross-Linguistic Evaluation of Code-Switching Constraints in Urban Multilingual Corpora

Adrian Thomas

Professor

Leiden University

Rapenburg 70, 2311 EJ Leiden, Netherlands

Morgan Harris

Associate Professor

Humboldt-Universität zu Berlin

Unter den Linden 6, 10099 Berlin, Germany

Quinn Walker

PhD

National University of Singapore

21 Lower Kent Ridge Road, Singapore 119077, Singapore

Nico Brown

D.Sc

University of Toronto

27 King's College Circle, Toronto, Ontario M5S 1A1, Canada

Abstract. This study investigates the competing theoretical frameworks of morphosyntactic convergence and pragmatic divergence as explanatory models for code-switching behaviour in urban multilingual communities. Drawing on a curated corpus of 1.2 million tokens collected across four metropolitan speech communities in Western Europe and Southeast Asia, we apply Optimality-Theoretic constraint ranking alongside Conversation Analysis protocols to evaluate the predictive validity of each approach. Quantitative corpus interrogation was supplemented by sociolinguistic interviews and discourse-completion tasks. Findings indicate that morphosyntactic convergence accounts adequately for intra-sentential switching patterns, while pragmatic divergence more robustly explains interactional alignment phenomena. Neither framework in isolation achieves comprehensive explanatory coverage, suggesting that an integrated, constraint-interaction model is required for empirically grounded analyses of contemporary multilingual discourse.

Keywords: code-switching, morphosyntactic convergence, pragmatic divergence, multilingual corpora, Optimality Theory, constraint ranking, intra-sentential switching, urban multilingualism, discourse-completion tasks

Introduction

The phenomenon of code-switching — the alternation between two or more languages or linguistic varieties within a single conversational episode or utterance — has occupied a central position in the sociolinguistic and formal linguistics literature for several decades. Yet, as urban multilingual communities continue to expand and diversify in the context of accelerating globalisation, digital communication, and transnational migration, the theoretical apparatus available for explaining the structural and pragmatic motivations underlying switching behaviour remains sharply contested. By 2024, metropolitan speech communities across Western Europe, Sub-Saharan Africa, and Southeast Asia exhibit contact configurations of unprecedented complexity, routinely involving three or more typologically distinct languages in everyday interaction. This sociolinguistic reality demands renewed theoretical scrutiny, since the binary opposition that has historically dominated the field — pitting structurally oriented frameworks against discourse-pragmatic accounts — appears insufficient to capture the full range of observed switching patterns. Two theoretical traditions have principally competed for explanatory primacy in this domain. The morphosyntactic convergence hypothesis, rooted in the Matrix Language Frame model (Myers-Scotton, 1993) and subsequently elaborated through Optimality-Theoretic constraint ranking (Prince & Smolensky, 2004), posits that lexical insertion and morphological accommodation during switching are governed by hierarchically ordered, language-internal structural constraints. Conversely, the pragmatic divergence framework — drawing on Gricean inferential pragmatics, Relevance Theory (Sperber & Wilson, 1986), and Conversation Analysis — contends that switching is primarily an interactional resource deployed for purposes of footing negotiation, affective stance marking, and audience design. These competing paradigms have generated divergent methodological commitments, with morphosyntactic researchers gravitating toward large-scale corpus annotation and formal constraint modelling, while pragmatic divergence scholars have favoured ethnographic and interactional data collection. Despite extensive individual elaboration of both frameworks, direct empirical comparison using matched corpora and unified analytical protocols is conspicuously absent from the literature. Furthermore, recent corpus-linguistic investigations leveraging computational annotation tools — including dependency parsing, Part-of-Speech tagging for mixed-language utterances, and automated interactional sequence detection — have created new methodological opportunities that neither tradition has fully exploited. The present study addresses this gap by subjecting a rigorously compiled, cross-linguistic multilingual corpus to parallel analytical procedures derived from both frameworks, thereby enabling a controlled, criterion-referenced comparison of their predictive validity. The remainder of this paper is structured as follows: Section 2 reviews the theoretical foundations and prior empirical literature pertaining to both morphosyntactic convergence and pragmatic divergence models. Section 3 details the corpus compilation methodology, annotation schema, and sociolinguistic interview protocol. Section 4 presents quantitative and qualitative results obtained through Optimality-Theoretic constraint ranking and Conversation Analysis respectively. Section 5 offers a comparative discussion of the two frameworks' explanatory scope, with particular

attention to intra-sentential switching and interactional alignment phenomena. Section 6 proposes an integrated constraint-interaction model reconciling the strongest predictive components of each approach. Section 7 concludes with implications for multilingual language policy, pedagogical grammar design, and future corpus-based research.

This is a preliminary version. To read the full version of the article, please purchase a subscription.

References

1. Gojayeveva, L. (2025). Modern methods of teaching literature. OEIL Research Journal, 23(11), 304.