University of Pannonia, Veszprém
Faculty of Humanities
Multilingualism Doctoral School

PhD Dissertation

Hungarian-English Code-Switching:

A comparative analysis of code-switching frameworks on morphosyntactic integration



Written by: Salamon Attila

Supervisor: Prof. Margaret Deuchar

Veszprém

2026

Abstract

Code-switching (CS) is generally understood as the use of more than one language in the same speech event, utterance, sentence or discourse. This study aims to demonstrate how naturally occurring Hungarian-English code-switching is produced, with special regard to morphosyntactic integration. The treatment of English-origin material in this language pair is compared under the Matrix Language Frame (MLF) Model and the Poplackian variationist approach, including the Nonce Borrowing Hypothesis (NBH). The study also aims to promote the use of publicly available podcast data for building linguistic corpus. The analysis draws on approximately 30 hours of naturally occurring speech from 24 Hungarian podcast videos involving 37 speakers. The corpus contains 419 bilingual constructions identified in 407 utterances, of which 405 intraclausal constructions were retained for structural comparison. The predictions of the two frameworks are tested based on the grammatical outcomes of the integration of English-origin material.

The results show extensive and systematic Hungarian grammatical control. Of the 405 intraclausal constructions, 187 involve intraword integration, including 152 lone English-origin items and 35 multiword expressions. English-origin bases received Hungarian case, number, possessive, verbal, adjectival, comparative, and nominalizing morphology and were placed in Hungarian coverb constructions. All 40 English-origin verbal constructions were formed by Hungarian verbalizer morphemes before possible further derivation or inflection.

The MLF Model provided an available structural analysis for all 405 intraclausal constructions: Hungarian consistently provided the morphosyntactic frame, and no unambiguous violation of the Morpheme Order Principle or the relevant system-morpheme predictions was identified. The Poplackian approach provided analysis for 370 constructions but left the 35 morphologically integrated multiword expressions indeterminate. As the frameworks mainly converge on the observable Hungarian structure while differing in the interpretation of integrated English-origin material as code-switching or borrowing, the conclusion that morphosyntactic integration can be established independently of contact-status classification is drawn. The study also highlights and demonstrates the value of podcasts as recheckable sources of researcher-unelicited bilingual speech, while recognizing their public, possibly performed, and non-representative characteristics.

Keywords: bilingualism, code-switching, intraclausal switching, intraword switching, morphosyntax, linguistic integration, podcast data, naturalistic speech, Matrix Language Frame Model, Nonce Borrowing Hypothesis, linguistic framework comparison

1. INTRODUCTION

Speakers of more than one language have always been in the focus of linguistic research; however, different methods, frameworks and language pairs might yield different results, therefore diverging conclusions can be drawn from such language-contact studies. This dissertation is written with the aim of testing different frameworks applied on the same corpus in an objective and scientifically sound manner, while analysing how other-language material is being incorporated during spontaneous speech.

Hungarian-English bilingual speech frequently combines English lexical material with Hungarian grammatical structure. In the podcast corpus examined in this dissertation, speakers produced forms such as *content-et* 'content-ACC', *stream-el-ünk* 'stream-VERB-1PL', and *fake meat-ek-ért* 'fake meat-PL-CAUS-FIN'. These examples contain recognizably English lexical material, yet their case marking, category assignment, number, inflection, and morpheme order are supplied by Hungarian.

Code-switching is generally understood as the use of more than one language within the same interaction, utterance, clause, or constituent (Bullock & Toribio, 2009). It is a systematic feature of bilingual speech rather than linguistic confusion or incomplete multilingual competence (Grosjean, 1989). Bilingual speakers may alternate between longer stretches of their languages, insert lexical elements, or construct forms in which material from both languages occurs within a single word (Gardner-Chloros, 2009; Poplack, 1980).

The present dissertation is concerned primarily with intraclausal and intraword combinations. The data are not extended alternations between complete Hungarian and English clauses, but constructions in which English-origin material participates in Hungarian syntax and morphology. The focus is less on why a speaker chooses an English expression in discourse and more on how that expression is grammatically incorporated.

According to the MLF Model, bilingual clauses are structurally asymmetrical. One language supplies the morphosyntactic frame, while the other contributes Embedded Language [EL] material. English lexical elements can receive Hungarian morphology and be analysed as code-switches. On this account, forms such as the examples of Table 1 demonstrate the control of the Hungarian Matrix Language [ML] over English Embedded Language material. This asymmetrical organization is central to this model (Myers-Scotton, 1993, 2002).

The Poplackian approach emphasises the distinction of code-switching from borrowing through grammatical behaviour. Lone donor-language [L^D] items that pattern with the recipient language [L^R] are analysed as established or nonce borrowings rather than as code-switches. While morphosyntactic integration is central to their classification, phonological integration is less decisive because these processes do not necessarily occur and develop together. The possibility that a socially unestablished item can immediately display recipient-language grammar is formalized in the Nonce Borrowing Hypothesis (Poplack, 2012, 2018; Sankoff et al., 1990).

Aims of the Dissertation

The primary aim of this dissertation is to describe how English-origin material is integrated into Hungarian morphosyntax in naturally occurring podcast speech. The second aim is to evaluate the extent to which the MLF Model accounts for the structural organization of the resulting bilingual constructions, with particular attention to ML identification, morpheme order, system morphology, category assignment, and the treatment of complex English material. Third, the study examines the same data from the Poplackian perspective, focusing on the distinction between integrated lone items, borrowing, nonce borrowing, and multiword code-switches. Finally, to compare the approaches and where they converge or diverge in their description of the data, classification, and which constructions remain theoretically indeterminate.

Research Questions and Predictions

Considering the aims, the dissertation addresses the following research questions:

1. How is English-origin material integrated into Hungarian morphosyntax in naturally occurring bilingual speech?
2. To what extent can the intraclausal constructions be accounted for by the Matrix Language Frame Model?
3. To what extent are the data compatible with the Poplackian approach and the Nonce Borrowing Hypothesis?
4. Where do the frameworks converge and diverge in their treatment of the same constructions?

2. CODE-SWITCHING AND BILINGUAL LANGUAGE PRODUCTION

Code-switching is one of the most noticeable manifestations of bilingual language use, yet its general description as the alternation between two languages covers considerable terminological and theoretical complexity. In its broadest sense, it refers to the use of two or more languages within the same interaction, discourse, sentence, or utterance. Research on bilingual speech has demonstrated that such alternation is not arbitrary or indicative of linguistic deficiency but is a systematic feature governed by grammatical, psycholinguistic, sociolinguistic, and interactional factors (Auer, 1998; Bullock & Toribio, 2009; Gardner-Chloros, 2009; Gumperz, 1982; Milroy & Muysken, 1995; Muysken, 2000). Consequently, contemporary approaches generally treat code-switching as part of bilingual linguistic competence. Code-mixing, language alternation and code-switching are often used as synonymous terms, although most of the literature discussed in this study applies the latter one.

Foundational research on bilingualism also warns against evaluating bilingual language use as the imperfect combination of two monolingual systems. Grosjean (1982, 1989) supports a holistic view according to which the bilingual is a specific and integrated speaker – hearer whose linguistic capabilities reflect the coexistence and interaction of more than one language. According to this view, bilingual performance should not be measured against the expectation that each language will behave exactly as it does in a monolingual speaker. From a Hungarian psycholinguistic perspective, Navracsics (2010) similarly treats code-switching and code-mixing in spontaneous speech as manifestations of bilingual language processing that may occur at several structural levels, including morpheme boundaries, individual lexical items, and larger phraseological units.

The discussion on the relationship between code-switching and borrowing has been in a central position in language-contact research since the foundational work of Haugen (1950) and Weinreich (1953). According to them, borrowing is generally associated with the incorporation of material originating in one language into another, however, determining when such incorporation takes place is much more difficult than identifying the historical origin of a lexical form. Frequency, conventionalization, phonological adaptation, morphological integration, syntactic behaviour, community distribution, and dictionary attestation have all been used as potential criteria. However, classification is difficult as theoretical approaches differ in the emphasis attributed to each of these criteria.

The production of bilingual speech requires speakers to manage linguistic knowledge associated with more than one language while constructing an utterance. Although the observable result may involve only one language or a combination of two languages, bilingual language production does not involve one linguistic system being fully activated while the other remains completely unavailable. Psycholinguistic research instead indicates that bilingual processing involves interaction between linguistic representations associated with both languages (Dijkstra & van Heuven, 2002; Green, 1998; Kroll & Stewart, 1994). Therefore, lexical activation, language selection, inhibition, proficiency, and interactional context are directly relevant to understanding how an English-origin lexical item can become available during predominantly Hungarian speech and enter a Hungarian morphosyntactic structure.

Models of bilingual production partly build on general theories of speech production. Levelt's (1989) influential work distinguishes between conceptual preparation, formulation, phonological encoding, and articulation. Although originally developed as an account of monolingual speech production, his model provides an important stepping stone for considering the additional demands faced by bilingual speakers. At the conceptual level, the speaker develops the preverbal message to be expressed, then retrieves lexical and grammatical information and organises it before the final phonological representation is articulated. In bilingual speakers, however, lexical knowledge is distributed across their two linguistic systems, where lexical candidates of different languages become activated and selected.

As natural bilingual language use can involve substantially different control demands from experimentally cued alternation, the bilingual production system must be considered in relation to the contexts in which it is used. Green and Abutalebi's (2013) Adaptive Control Hypothesis develops this perspective by proposing that language-control processes adapt to the interactional demands. Different bilingual environments impose different requirements on language selection. Thus, the production of an English-origin item during an otherwise Hungarian discourse should not be interpreted as a failure of lexical retrieval or language control automatically. Han et al. (2022) found associations between habitual switching patterns and performance in verbal switching, nonverbal shifting, and inhibition tasks, which provides support for the Adaptive Control Hypothesis, and similarly, Özkara et al. (2025) emphasize that conclusions about code-switching and cognitive control depend on the operationalisation of switching experience, interactional context, and control processes.

A substantial part of specifically Hungarian-English code-switching literature has been based on Hungarian-American bilingual communities and has approached switching from sociolinguistic, pragmatic, grammatical, or cognitive perspectives. Myers-Scotton and Bolonyai (2001) analyse code-switching within the Rational Choice Model. They argue that linguistic choices can be understood as individually made but socially constrained decisions where speakers select among the available linguistic alternatives based on their communicative goals, costs and benefits. Combining Hungarian-English data with theoretically explicit analyses of bilingual grammar, Bolonyai's (2005) study *English Verbs in Hungarian/English Codeswitching*, investigates the grammatical integration of English verbs into Hungarian. Her focus is on why Hungarian employs morphological integration when incorporating English verbs.

Differentiating and classifying cases of borrowings and code-switches has been attempted by Bolonyai (2005) by referring to morphosyntactically integrated English forms that retain English-like pronunciation as CS, while forms displaying both morphosyntactic and phonological integration as borrowings for the purposes of her study. This suggests that even within the Hungarian-English literature, the same observable property, i.e. integration, does not necessarily point to the same classification of the English-origin material.

3. MORPHOSYNTACTIC INTEGRATION

Morphosyntactic integration refers to the extent to which linguistic material originating in one language conforms to the morphological and syntactic requirements of another language when it is incorporated into bilingual discourse. In contact-linguistic research, integration is therefore not limited to the attachment of overt morphology (Deuchar & Stammers, 2016; Poplack, 2018). It may involve several related processes, including the assignment of grammatical category, the use of recipient- or Matrix-Language derivational morphemes, inflections, and the accommodation of the item to language-specific word-order requirements. Thus, other-language elements can be integrated at more than one grammatical level and to different degrees.

This phenomenon has played an important role in attempts to distinguish code-switching from borrowing. Earlier contact-linguistic work already recognized that lexical material from a language may undergo structural adaptation upon entering another one (Haugen, 1950; Weinreich, 1953). Later variationist research made the grammatical behaviour of other-

language material the main empirical criterion. The works of Poplack and Sankoff (1984), Poplack et al. (1988), and Poplack (2018) emphasize that the relevant question is not simply whether an item carries an overt recipient-language affix, but whether it patterns morphosyntactically the same way as the comparable recipient-language lexical items. Distinguishing the two is important because recipient-language grammatical marking itself can be variable, and the absence of overt affixation does not necessarily imply the absence of integration if the native forms exhibit the same variation.

The grammatical interpretation of morphosyntactic integration may vary according to theoretical framework. In the MLF Model, morphosyntactic asymmetry is a basic property of most mixed constituents where the Matrix Language supplies the grammar into which Embedded Language material can be incorporated under specified structural conditions (Myers-Scotton, 1993, 2002). On the other hand, the Poplackian approach requires morphosyntactic patterning to determine the borrowing status of the donor-language item. Also, the findings of Budzhak-Jones (1998) on her analysis of Ukrainian-English bilingual discourse illustrate that an intraword language boundary is an observable structural property, but its classification as code-switching cannot be assumed solely from the different lexical origins of the stem and the affix, so she applied the term borrowing to such cases.

Since Hungarian is commonly characterized as a morphologically rich, predominantly agglutinative language, with extensive affixation in both nominal and verbal paradigms (É. Kiss, 2002; Kenesei et al., 1998; Kiefer, 2000), it is particularly suitable for investigating morphosyntactic integration; even in literary works, where such integration might carry additional functions and communicative goals, resulting in further classification (Domokos & Deganutti, 2021; T. Kovács, 2024).

Typological distance between English and Hungarian has direct consequences for bilingual constructions: an English lexical item inserted into Hungarian discourse can enter environments where Hungarian requires overt realization of grammatical properties that English either expresses analytically or does not mark on the corresponding item.

Constituent size, internal grammatical organization, external Hungarian integration, and morphological complexity are considered as somewhat independent dimensions. Isolated items may be extensively integrated morphologically, while multiword English expressions may retain their internal structure and receive Hungarian marking. Therefore, the dissertation distinguishes among LOLIs, intraword mixed forms, mixed constituents, mixed compounds,

and internally English multiword constituents. These descriptive categories provide a common basis for the framework-specific analyses in Chapter 4 without assuming in advance whether any individual construction is a borrowing, code-switching, insertion, alternation, or an Embedded Language island.

4. COMPETING ACCOUNTS OF BILINGUAL MORPHOSYNTAX

The Matrix Language Frame Model

The Matrix Language Frame (MLF) Model is one of the most influential accounts of the morphosyntactic organization of intrasentential code-switching. Developed and later refined by Myers-Scotton (1993, 2002), the model proposes asymmetrical contribution to bilingual grammatical structure: one language provides morphosyntactic frame called the Matrix Language, while material from another is embedded.

Since English-origin lexical material frequently occurs with Hungarian derivational and inflectional morphology and within Hungarian syntactic structure, the framework is particularly relevant to the present data. The following sections outline the components of the model that generate testable predictions for such bilingual constructions.

The original MLF Model was developed to explain recurrent structural asymmetries in intrasentential code-switching and not necessarily to specify possible switch points (Myers-Scotton, 1993). The distinction between the ML and the EL was formulated primarily through the Morpheme Order Principle and the System Morpheme Principle. This was later expanded by the 4-M Model, which provides a more detailed classification of morphemes. The Abstract Level Model, concerning different levels of lexical structure, the Uniform Structure Principle and the Asymmetry Principle are further developments of the MLF and should not be treated as components of a single unchanged model (Myers-Scotton & Jake, 2000; Myers-Scotton, 2002).

Although the MLF Model provides testable predictions for bilingual morphosyntax, several aspects of its application require caution. One methodological issue concerns the identification of the Matrix Language. If the same morphosyntactic properties are used both to establish ML status and then to confirm Matrix Language control, the analysis might risk circularity. For this

reason, ML identification and the testing of individual MLF predictions should therefore be kept analytically independent if possible.

The Poplackian Variationist Approach

Poplack's approach combines grammatical analysis with variationist methodology to distinguish different forms of bilingual language mixing empirically. Its early development focused on grammatical constraints on intrasentential code-switching, namely the Equivalence and Free Morpheme Constraints (Poplack, 1980; Sankoff & Poplack, 1981). Her and her colleagues' later work increasingly focused on the distinction between code-switching and lexical borrowing, especially the status of LOLIs occurring within recipient-language discourse (Poplack & Sankoff, 1984; Poplack et al., 1988, 1989; Sankoff et al., 1990; Poplack, 2012, 2018).

Although the Poplackian view approaches these structures through different diagnostics, the same empirical problem is examined: how English-origin material behaves when incorporated into Hungarian morphosyntax. This approach, however, places particular emphasis on whether donor-language-origin material patterns grammatically with the recipient language and whether its distribution is characteristic of borrowing or code-switching.

The strength of the approach in discussion also poses a rather big methodological demand. Identifying a form as indeed nonce requires evidence of its recurrence and community distribution, while the establishment of L^R-equivalent behaviour requires sufficient comparable native, borrowed, and switched material.

This is quite important for the present podcast corpus: recurrence across speakers, external corpus frequency, and dictionary attestation can provide useful evidence concerning the establishment of the items, but they are not identical to the community-distribution evidence which is a feature of the original variationist methodology. Also, an item occurring once in the corpus cannot automatically be assumed to be nonce, and widespread dictionary or corpus attestation need to be treated as evidence of establishment and not of its sociolinguistic status in the sampled community.

5. DATA AND METHODS

The present study adopts a corpus-based, analytical design to investigate morphosyntactic integration in naturalistic Hungarian-English bilingual speech. The data consist of publicly available Hungarian-language podcast recordings produced independently of the research. After their initial selection, the bilingual utterances, treated as units of analysis, were manually extracted from the podcasts recordings and were transcribed and analysed in both frameworks.

First, English-origin material is coded according to observable structural properties without assigning an immediate code-switching or borrowing classification. These properties include constituent size, word class, Hungarian derivational and inflectional morphology, case and possessive marking, verbalization, coverb use, constituent structure, and the internal organization of multiword English material. This theory-neutral coding provides a common empirical basis for the later framework-specific analyses. Later, the coded constructions are evaluated separately under the MLF Model and the Poplackian approach and their predictions introduced in the previous chapter.

Framework-level coverage and prediction-specific evidence are handled separately. Individual predictions are tested only in constructions that provide the grammatical or distributional evidence required for that test; non-diagnostic structures are not counted as equal confirmations. At the broader framework level, constructions may receive a determinate analysis, conflict with an applicable prediction, or remain indeterminate where the framework or available information does not permit a reliable analysis. This design makes it possible to compare both the empirical support for individual predictions and the analytical coverage of the two frameworks while keeping overt morphosyntactic integration separate from a specific interpretation.

The dataset consists of naturally occurring Hungarian-English bilingual speech collected from publicly available Hungarian-language podcasts uploaded to YouTube. It contains approximately 30 hours of video-recorded material from 24 podcast episodes and includes 37 speakers between 20 and 69 years of age (at the time of recording). In all cases, Hungarian is the first and dominant language of the speakers. Based on their observable use of English in the recordings, their proficiency was estimated to range broadly from A2 to C1 level; no standardized proficiency assessment was available due to the nature of the data. These recordings represent mostly informal and semi-formal conversational settings.

A dictionary criterion was used to exclude clearly conventionalised Hungarian loans from the dataset. English-origin forms listed as established Hungarian lexical items in *Magyar értelmező kéziszótár* [*Concise Explanatory Dictionary of Hungarian*] (Akadémiai Kiadó, n.d.) were excluded. Forms not listed in the reference source were provisionally retained where their English origin was otherwise identifiable. This procedure follows the operational use of dictionary evidence in previous corpus-based code-switching research (Deuchar, 2006).

The initial dataset included a total of 419 English-origin occurrences within 407 bilingual utterances. Of these, 405 were intraclausal and 14 interclausal. Only the intraclausal occurrences were retained for the later morphosyntactic analyses.

The 405 intraclausal constructions consisted of 302 lone or single-item insertions and 103 multiword constructions, representing 74.6% and 25.4% of the intraclausal dataset, respectively. Within the intraclausal set, a total of 187 constructions contained an intraword language boundary as well.

It is important to note that intraword status was treated as an overlapping structural property rather than as a separate constituent-size category, since of these constructions, 152 involved LOLIs and 35 involved multiword English material carrying Hungarian morphology. The remaining dataset consisted of 150 lone items and 68 multiword constructions without intraword morphological integration.

6. RESULTS

As stated above, the corpus contained 407 bilingual utterances, within which 419 individual English-origin occurrences were identified. After excluding the 14 occurrences with interclausal switches the resulting analytical dataset consisted of 405 intraclausal constructions occurring within 393 bilingual utterances.

Of these, 302 (74.6%) were lone or single-item constructions and 103 (25.4%) were multiword constructions. English-origin material occurred as nouns, adjectives, verbal bases, adverbs, discourse markers, compounds, fixed expressions, and larger nominal or other sequences.

Intraword integration occurred in 187 constructions (46.2%). Of these, 152 were lone items and 35 were multiword constructions carrying Hungarian morphology. The remaining dataset consisted of 150 lone items without intraword morphology and 68 multiword constructions

without intraword morphology. The four resulting structural categories, based on intraword integration and the number of English words of the insertions, were the following:

- lone, non-intraword constructions: 150 (37.0%);
- lone intraword constructions: 152 (37.5%);
- multiword, non-intraword constructions: 68 (16.8%);
- multiword intraword constructions: 35 (8.6%).

Thus, constituent size and intraword integration were treated as independent dimensions. While most constructions contained only one English-origin lexical item, Hungarian morphology also occurred on 35 of the 103 multiword constructions (34.0%).

Following the framework-specific analyses, the overall analytical outcomes were compared across the 405 intraclausal constructions. As defined in Chapter 5, constructions were classified at framework level as consistent, inconsistent, or indeterminate. The table below summarises the coverage of the two frameworks on all intraclausal cases.

Framework	**Consistent cases**	**Inconsistent cases**	**Indeterminate cases**	**Total**
MLF	405 (100.0%)	0 (0.0%)	0 (0.0%)	405
Poplackian	370 (91.4%)	0 (0.0%)	35 (8.6%)	405

Conclusively, under the MLF analysis, all 405 intraclausal constructions received a determinate structural analysis. Hungarian was identified as the Matrix Language throughout the dataset, and no construction showed an unambiguous violation of the applicable predictions.

7. DISCUSSION

The findings show extensive and systematic integration of English-origin material into Hungarian morphosyntax. Of the 419 English-origin occurrences identified in the corpus, 405 were intraclausal and formed the basis of the structural comparison. These included 302 lone English-origin items and 103 multiword constructions. Intraword integration occurred in 187 cases, comprising 152 lone items and 35 multiword constructions, where English-origin material participated in Hungarian case, number and possessive marking, verbal and adjectival derivation, degree formation, nominalization, verbal inflection, and coverb constructions. Several forms contained multiple successive layers of Hungarian derivational and inflectional morphology.

This structural pattern answers the first research question clearly: English-origin material was not merely inserted alongside Hungarian material but regularly entered grammatical structures organized by Hungarian. Particularly revealing examples included verbal formations such as *stream-el-ünk*, multilayered forms such as *gameplay-ez-és-nél* and *feature-e-i-t*, and externally inflected multiword expressions such as *fake meat-ek-ért* and *community guideline set-tel*; however, constituent size and morphological integration did not vary categorically together, as Hungarian morphology could apply to both lone lexical items and to complex English-origin bases.

The MLF analysis provided a determinate structural account of all 405 intraclausal constructions, with Hungarian being identifiable as the Matrix Language throughout the corpus. Diagnostic intraword constructions followed Hungarian morpheme order, and the relevant outsider late system morphemes came from Hungarian. Also, English-origin lexical material conformed to Hungarian category-forming and inflectional requirements, and internally English sequences could be integrated through the distinction between ML+EL constituents and Embedded Language islands. In sum, no unambiguous violation of the applicable of the model's predictions was identified.

The Poplackian framework also accounted for most structural tendencies, with 152 overtly integrated lone items displaying Hungarian morphosyntax, while the clearest internally structured multiword sequences retained English grammar. These findings were compatible with the distinction between L^R-integrated borrowing and L^D-structured multiword code-switching.

Evidence for the NBH was more limited: apparently non-established forms could display immediate Hungarian integration, which is predicted by the model, but their nonce status could not normally be established independently. What is more, 35 morphologically integrated multiword constructions remained indeterminate under the canonical Poplackian analysis.

The most important comparative result is probably that the grammatical structure of a bilingual construction can be considerably clearer than its classification as CS or borrowing. While both frameworks recognize the Hungarian contribution to the external structure of the data, they do not assign the same theoretical status to the English-origin material.

8. CONCLUSIONS

This dissertation examined the morphosyntactic integration of English-origin material in naturally occurring Hungarian podcast speech and compared the same constructions under the Matrix Language Frame (MLF) Model and the Poplackian variationist approach, including the Nonce Borrowing Hypothesis (NBH). The corpus contained 419 bilingual forms, of which 405 intraclausal constructions were retained for structural analysis.

The findings show a highly consistent pattern of Hungarian grammatical organization. English supplied primarily lexical material, whereas Hungarian provided category formation, inflection, case, possession, verbal morphology, and external constituent structure. At the same time, the analysis demonstrated that grammatical integration can often be established more confidently than the resulting form can be classified as code-switching or borrowing.

Answers to the Research Questions

The first research question asked how English-origin material is integrated into Hungarian morphosyntax. As presented through the data, integration occurred across nominal, verbal, and adjectival domains and included case, number, possession, verbalization, adjectival derivation, comparison, nominalization, verbal inflection, and coverb constructions. Of the 405 intraclausal constructions, 187 contained intraword integration, including 152 lone items and 35 morphologically integrated multiword constructions. All 40 English-origin verbal constructions contained Hungarian verbalizing morphology, confirming and extending the integration pattern described by Bolonyai (2005).

The second research question concerned the MLF Model. In its analysis, Hungarian was identifiable as the Matrix Language throughout the intraclausal dataset, and no unambiguous

violation of the applicable Morpheme Order or System Morpheme predictions was identified. Diagnostic intraword constructions followed Hungarian morpheme order, the relevant outsider late system morphemes originated in Hungarian, and internally English sequences could be accommodated as EL islands where English internal grammar was independently identifiable. The MLF Model provided a determinate structural analysis for all 405 constructions, although not every construction was equally diagnostic.

The third research question concerned the Poplackian approach and the NBH. The 152 overtly integrated lone English-origin items displayed clear Hungarian morphosyntax, while the clearest multiword English sequences retained English internal structure. Therefore, the data were compatible with the Poplackian distinction between recipient-language-integrated borrowing and donor-language-structured switching (Poplack, 2018). They were also compatible with the NBH prediction that grammatical integration does not require prior social establishment (Sankoff et al., 1990; Poplack, 2012; Poplack & Dion, 2012). However, nonce status could not be demonstrated independently for individual items, and no complete matched variationist comparison with native Hungarian forms was conducted. Also, the Poplackian analysis provided a determinate account for 370 constructions, while 35 morphologically integrated multiword constructions remained indeterminate.

The fourth research question asked where the frameworks converge and diverge. Both approaches agreed that Hungarian supplied the relevant external grammatical structure and that clearly structured English multiword sequences could preserve English internal grammar. Their main disagreement concerned the interpretation of integrated English-origin material. Under the MLF, such material may remain EL material within a Hungarian ML frame; under the Poplackian approach, L^R-equivalent grammatical behaviour is borrowing-compatible. The difference between 100% MLF coverage and 91.4% Poplackian coverage should be interpreted as a difference in their success in analytical coverage.

References

Akadémiai Kiadó. (n.d.). *Magyar értelmező kéziszótár [Concise explanatory dictionary of Hungarian]*. szotar.net. Retrieved February 3, 2026, from https://www.szotar.net/

Auer, P. (Ed.). (1998). *Code-switching in conversation: Language, interaction and identity.* Routledge.

Bolonyai, Á. (2005). English verbs in Hungarian/English codeswitching. In J. Cohen, K. T. McAlister, K. Rolstad, & J. MacSwan (Eds.), *Proceedings of the 4th International Symposium on Bilingualism* (pp. 317–327). Cascadilla Press.

Budzhak-Jones, S. (1998). Against word-internal codeswitching: Evidence from Ukrainian–English bilingualism. *International Journal of Bilingualism, 2*(2), 161–182. https://doi.org/10.1177/136700699800200203

Bullock, B. E., & Toribio, A. J. (Eds.). (2009). *The Cambridge handbook of linguistic code-switching.* Cambridge University Press.

Deuchar, M. (2006). Welsh–English code-switching and the Matrix Language Frame model. *Lingua, 116*(11), 1986–2011. https://doi.org/10.1016/j.lingua.2005.06.001

Deuchar, M., & Stammers, J. R. (2016). English-origin verbs in Welsh: Adjudicating between two theoretical approaches. *Languages, 1*(1), Article 7. https://doi.org/10.3390/languages1010007

Dijkstra, T., & van Heuven, W. J. B. (2002). The architecture of the bilingual word recognition system: From identification to decision. *Bilingualism: Language and Cognition, 5*(3), 175–197. https://doi.org/10.1017/S1366728902003012

Domokos, J., & Deganutti, M. (2021). Four major literary code-switching strategies in Hungarian literature: Decoding monolingualism. *Hungarian Studies Yearbook, 3*(1), 43–63. https://doi.org/10.2478/hsy-2021-0004

É. Kiss, K. (2002). *The syntax of Hungarian.* Cambridge University Press.

Gardner-Chloros, P. (2009). *Code-switching.* Cambridge University Press.

Green, D. W. (1998). Mental control of the bilingual lexico-semantic system. *Bilingualism: Language and Cognition, 1*(2), 67–81. https://doi.org/10.1017/S1366728998000133

Green, D. W., & Abutalebi, J. (2013). Language control in bilinguals: The adaptive control hypothesis. *Journal of Cognitive Psychology, 25*(5), 515–530. https://doi.org/10.1080/20445911.2013.796377

Grosjean, F. (1982). *Life with two languages: An introduction to bilingualism.* Harvard University Press.

Grosjean, F. (1989). Neurolinguists, beware! The bilingual is not two monolinguals in one person. *Brain and Language, 36*(1), 3–15. https://doi.org/10.1016/0093-934X(89)90048-5

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

Han, X., Li, W., & Filippi, R. (2022). The effects of habitual code-switching in bilingual language production on cognitive control. *Bilingualism: Language and Cognition, 25*(5), 869–889. https://doi.org/10.1017/S1366728922000244

Haugen, E. (1950). The analysis of linguistic borrowing. *Language, 26*(2), 210–231. https://doi.org/10.2307/410058

Kenesei, I., Vago, R. M., & Fenyvesi, A. (1998). *Hungarian.* Routledge.

Kiefer, F. (Ed.). (2000). *Strukturális magyar nyelvtan 3: Morfológia [Structural grammar of Hungarian 3: Morphology].* Akadémiai Kiadó.

Kovács, T. (2024). The meaning of code-switching in Anne Tardos's multilingual poem *Minden Ami*. *Orpheus Noster, 16*(3), 12–23.

Kroll, J. F., & Stewart, E. (1994). Category interference in translation and picture naming: Evidence for asymmetric connections between bilingual memory representations. *Journal of Memory and Language, 33*(2), 149–174. https://doi.org/10.1006/jmla.1994.1008

Levelt, W. J. M. (1989). *Speaking: From intention to articulation.* MIT Press.

Milroy, L., & Muysken, P. (Eds.). (1995). *One speaker, two languages: Cross-disciplinary perspectives on code-switching.* Cambridge University Press.

Muysken, P. (2000). *Bilingual speech: A typology of code-mixing.* Cambridge University Press.

Myers-Scotton, C. (1993). *Duelling languages: Grammatical structure in codeswitching.* Clarendon Press.

Myers-Scotton, C. (2002). *Contact linguistics: Bilingual encounters and grammatical outcomes.* Oxford University Press. https://doi.org/10.1093/acprof:oso/9780198299530.001.0001

Myers-Scotton, C., & Bolonyai, A. (2001). Calculating speakers: Codeswitching in a rational choice model. *Language in Society, 30*(1), 1–28. https://doi.org/10.1017/S0047404501001014

Myers-Scotton, C., & Jake, J. L. (2000). Four types of morpheme: Evidence from aphasia, code switching, and second-language acquisition. *Linguistics, 38*(6), 1053–1100. https://doi.org/10.1515/ling.2000.021

Navracsics, J. (2010). Kódváltás és kódkeverés kétnyelvűek spontán beszédében [Code-switching and code-mixing in bilinguals' spontaneous speech]. In J. Navracsics (Ed.), *Nyelv, beszéd, írás: Pszicholingvisztikai tanulmányok I [Language, speech, writing: Psycholinguistic studies I]* (pp. 121–129). Tinta Könyvkiadó.

Özkara, B., Cedden, G., von Stutterheim, C., & Meyer, P. (2025). Code-switching and cognitive control: A review of current trends and future directions. *Frontiers in Language Sciences, 4*, Article 1515283. https://doi.org/10.3389/flang.2025.1515283

Poplack, S. (1980). Sometimes I'll start a sentence in Spanish y termino en español: Toward a typology of code-switching. *Linguistics, 18*(7–8), 581–618. https://doi.org/10.1515/ling.1980.18.7-8.581

Poplack, S. (2012). What does the Nonce Borrowing Hypothesis hypothesize? *Bilingualism: Language and Cognition, 15*(3), 644–648. https://doi.org/10.1017/S1366728911000496

Poplack, S. (2018). *Borrowing: Loanwords in the speech community and in the grammar.* Oxford University Press.

Poplack, S., & Dion, N. (2012). Myths and facts about loanword development. *Language Variation and Change, 24*(3), 279–315. https://doi.org/10.1017/S095439451200018X

Poplack, S., & Sankoff, D. (1984). Borrowing: The synchrony of integration. *Linguistics, 22*(1), 99–135. https://doi.org/10.1515/ling.1984.22.1.99

Poplack, S., Sankoff, D., & Miller, C. (1988). The social correlates and linguistic processes of lexical borrowing and assimilation. *Linguistics, 26*(1), 47–104. https://doi.org/10.1515/ling.1988.26.1.47

Poplack, S., Wheeler, S., & Westwood, A. (1989). Distinguishing language contact phenomena: Evidence from Finnish–English bilingualism. *World Englishes, 8*(3), 389–406. https://doi.org/10.1111/j.1467-971X.1989.tb00677.x

Sankoff, D., & Poplack, S. (1981). A formal grammar for code-switching. *Papers in Linguistics, 14*(1), 3–45. https://doi.org/10.1080/08351818109370523

Sankoff, D., Poplack, S., & Vanniarajan, S. (1990). The case of the nonce loan in Tamil. *Language Variation and Change, 2*(1), 71–101. https://doi.org/10.1017/S0954394500000272

Weinreich, U. (1953). *Languages in contact: Findings and problems.* Linguistic Circle of New York.