Оригинални научни рад
811.163.41'373.611
https://doi.org/10.46793/NasKg2663.255T

Наслеђе 63 • 2026 • 255–274

Gorica R. Tomić[1]
University of Kragujevac
Faculty of Education in Užice
ORCID: https://orcid.org/0000-0002-8546-4781

RECONSTRUCTING NEW SERBIAN LEXICAL BLENDS: (SOCIO)LINGUISTIC EVIDENCE

Lexical blends are unconventional products created by combining (parts of) two or more source words, often involving segment loss and structural overlap, which makes their constituent elements difficult to recognize. This structural opacity poses challenges for speakers attempting to reconstruct the source words, a topic that has been only sparsely investigated in Serbian, particularly with regard to speaker comprehension. The aim of the present paper is twofold: (a) to investigate whether relative morphotactic transparency and overlapping contribute to native speakers' success in reconstructing Serbian blends and (b) to investigate the influence of several sociodemographic factors (sex, age, education level, occupation, and area of residence, specifically rural vs. urban) on the informants' success in blend reconstruction. The results of the quantitative analyses of the informants' correct answers show that complete-telescope blends are most successfully reconstructed, whereas the reconstruction of complete-inclusive blends is found to be least successful. It has also been shown that overlapping blends are more successfully reconstructed than non-overlapping ones. The results of the sociolinguistic analysis show that none of the four sociodemographic factors (i.e., sex, age, education level, and area of residence) has significant effect on the informants' success in recognizing the source words in a blend, but that there are statistically significant differences in the average number of correct answers between certain occupational groups, possibly due to varying degrees of linguistic creativity and exposure to language innovation in these occupational settings.

Keywords: *lexical blends*, *Serbian*, *native speakers*, *blend reconstruction*.

1 Introduction

In comparison to English, lexical blending in Serbian is a relatively recent, yet remarkably productive phenomenon.[2] The popularization of blending in Serbian in the late 1990s is believed to have been influenced by a number of factors, such as (linguistic) globalization, internationalization, as well as democratization of the Serbian language and society (Bugarski 2001: 1; 2016;

1 tomic@pfu.kg.ac.rs

2 Over the last 24 years, about 2,500 examples of Serbian blends have been recorded (see Bugarski 2001; 2019; Tomić 2019; 2023).

2019: 17; Halupka-Rešetar and Lalić-Krstin 2009: 115). We are nowadays witnessing the birth of new, mostly one-off, blends in Serbian practically every day and, perhaps more importantly, in diverse sociolinguistic registers.

Lexical blending is usually defined as a word-formation process in which two (or sporadically three) source words (SWs) and/or their splinters (i.e., submorphemic elements splintered off from SWs (Adams 1973; Fandrych 2008)) are consciously and intentionally combined into a single word, the meaning of which is a specific combination of (parts of) the SWs' meanings (cf. Bauer 2012: 11; Beliaeva 2019: n.p.; Renner 2023: 2–4). Examples from Serbian include *Bibiriki* 'name of a snack food' ← bib(er) 'pepper' × (kik)iriki 'peanut', *Testival* 'name of festival of dough, pasta, and bread' ← test(o) 'dough' × (f)estival 'festival', *vakcionalnost* 'used satirically to refer to a new type of national identity based on the Covid-19 vaccine one received' ← vakci(na) 'vaccine' × (n)acionalnost 'nationality' (Tomić 2023).[3]

As evidenced by the above examples from Serbian, blends are rather unconventional lexical formations. Because of their unconventionality as regards word-formation rules, blends seem to be most satisfactorily explained by Dressler and Merlini Barbaressi's (1994, as cited in Mattiello 2013: 1) term *Extra-grammatical Morphology* referring to "a set of heterogeneous formations (of an analogical or rule-like nature) which do not belong to morphological grammar, in that the processes through which they are obtained are not clearly identifiable and their input does not allow a prediction of a regular output" (Mattiello 2013: 1, cf. also Dressler 2000: 2), as is the case with the products of regular word-formation processes such as compounding or derivation. The fact that SWs suffer the loss of some of their segments, either through proper or improper clipping (i.e., overlapping) (cf. Mattiello 2013: 121), and that they are "not [therefore] transparently analysable into morphemes" (Mattiello 2022: 31) is probably one of the reasons why the SWs of blends are generally difficult to recognize (Mattiello 2022: 31).

Considering their structural transgression (Renner 2015), it is not surprising that blends, including Serbian examples, have most extensively been studied from a formal perspective (Bugarski 2019; Halupka-Rešetar and Lalić-Krstin 2009; Lalić-Krstin & Halupka-Rešetar 2007; Tomić 2019; 2022). Semantic aspects of Serbian blends have been studied much less extensively, possibly because the meanings of blends in general are frequently reduced to simple combinations of source-word meanings (Beliaeva 2019: n.p.; Bugarski 2001: 1; 2019: 17), despite the fact that they often involve additional evaluative, playful, or context-dependent nuances that are more difficult to analyze systematically (Bauer 2017: 159–160). It is, however, also necessary that we have insights into

3 In the examples, brackets indicate segments of the source words that do not participate in the blend formation. The symbol × marks the blending process between the source words. Overlapping material between SWs is indicated by underlining.

the ability of speakers to understand blends. Papers investigating the understanding of blends in Serbian are virtually non-existent, excluding the paper by Halupka-Rešetar & Lalić-Krstin (2012), who analyzed the influence of students' age, education level, and knowledge of English on the reconstruction of 30 Serbian blends, both in and out of context. Their results showed that only a very small percentage of the students were able to correctly identify both SWs. The students were most successful at reconstructing blends which exhibit a considerable degree of overlapping (e.g., *krađanin* ← krađa × (g)rađanin or *čedovište* ← čedo × č(u)dovište), in which SWs lose only one or two graphemes. On the other hand, examples that were most difficult to reconstruct were non-overlapping blends with a massive loss of SW material. Their results also indicated a positive correlation between the factors of age and education level and the students' ability to reconstruct the SWs in blends.

Sociolinguistic and psycholinguistic studies (e.g., Lehrer 1996, 2003; Halupka-Rešetar & Lalić-Krstin 2012; Silaški & Đurović 2013) have shown that linguistic processing and word-interpretation abilities may vary across sociodemographic groups, although results are often inconclusive and context-dependent. This suggests that such factors may also play a role in the reconstruction of Serbian blends.

By using a selection of new Serbian blends, our study aims to partly fill a clear gap in the literature – namely, the near absence of research on the reconstruction of Serbian blends, with only one study addressing this issue – by investigating the following research questions (RQs).

(1) Does morphotactic transparency of blends facilitate source-word recognizability?[4]

(2) Does overlapping, as a distinctive feature of blends, facilitate source-word recognizability?

(3) How do different sociodemographic factors (sex, age, education level, occupation, and area of residence, specifically rural vs. urban) influence the informants' success in blend reconstruction?

Based on the above research questions, three hypotheses can be formulated:

Hypothesis 1: Considering the fact that there is a general preference for morphotactic transparency in word-formation (Dressler 2005: 272), understood here as the degree of segmentability and perceptual accessibility of the constituent elements, (see Dressler 2005; Ronneberger-Sibold 2006), we hypothesize that the greater the morphotactic transparency of blends is – operationalized here following Ronneberger-Sibold's (2006) scale of complete-telescope (hereafter *telescope blends*), complete-inclusive (hereafter *inclusive*

4 Morphotactic transparency is understood here as the degree of segmentability and perceptual accessibility of the constituent elements (see Dressler 2005; Ronneberger-Sibold 2006). This notion is discussed in more detail in Subsection 2.1.

blends), contour, semi-complete, and fragment blends, supplemented by the technique of complete-superimposed blending (hereafter *superimposed blends*) introduced in Tomić (2023) – the more successful the informants will be at reconstructing the source words.

Hypothesis 2: Considering that overlaps are said to increase blend transparency (Ronneberger-Sibold 2006: 168–169), it is hypothesized that overlapping blends will be more successfully reconstructed than non-overlapping ones.

Hypothesis 3: Considering that most people nowadays have Internet access, it is hypothesized that sociodemographic factors (sex, age, education level, and area of residence) will not have a significant influence on informants' success in blend reconstruction, whereas differences between occupational groups may be statistically significant, possibly due to varying degrees of linguistic creativity and exposure to language innovation in their work. The paper is structured as follows: after the introduction, Section 2 describes the materials and methods used for (statistically) analyzing the data. Section 3 includes the (statistical) analysis and discussion of the results in light of the above hypotheses and results of similar research papers. Finally, Section 4 provides the most important conclusions as well as some implications for future (socio)linguistic research into Serbian blends.

2 Materials and methods

2.1 Blends

For the purposes of the present research, a subset of 50 blends was selected from a collection of 400 new Serbian blends presented in Tomić (2023), where "new" refers to blends first attested between 2002 and 2022, which were largely undocumented in the literature at the time of compilation, with two items lacking precise attestation data. The blends were selected so that all six blending types were represented in the dataset, although not in equal numbers, in order to enable comparison across types. Within each type, blends were selected non-systematically from the available collection, with preference given to examples that were more recently attested in usage. The blends are formed by six different blending techniques, given from most to least transparent, which are briefly explained and exemplified in the rest of this subsection. In line with Dressler (2005), morphotactic transparency is understood as a gradient property of word-formation structures, ranging from highly transparent formations with clearly segmentable constituents to opaque formations in which internal structure is not readily recoverable. Five of these techniques (i.e., telescope, inclusive, contour, semi-complete, and fragment blending) were introduced by Ronneberger-Sibold (2006: 168–169) as part of her typology based on the

degree of relative morphotactic and morphosemantic transparency of blends, while the sixth (complete-superimposed blending) was added in Tomić (2023).[5]

The blends in the sample are classified into six types, given from most to least transparent:

1) Telescope blends (*dnovinarka, tviteroristkinja, hranac*, *trološ*, *tviterapeut*, *zloslobođenje*, *kondukterijer*, *muzikaljenje*, *sarmada*, *grokenrol*), in which the end of SW1 continuously overlaps with the beginning of SW2 (Ronneberger-Sibold 2006: 168) (*hranac* ← hrana × ranac).[6]
2) Inclusive blends (*doČEPaj se*, *NINdže*, *nazaDNO*, *BORba, GIMnastičar*), in which one SW is included in the other, with which it overlaps (*doČEPaj se* ← dočepaj se × čep). These blends are "recognizable by spelling only" (Ronneberger-Sibold 2006: 168), which is why some authors refer to them as *graphic blends* (Bugarski 2019: 120). The included word is usually signaled by some graphic means, such as capital letters, special font type, etc.
3) Superimposed blends (*Pivonirka*, *Miškomor*, *zlaova*, *Svetlozar*), in which one full SW is placed over the other full SW with which it overlaps (usually discontinuously) (*Pivonirka* ← pivo × pionirka) (Tomić 2023: 104).
4) Contour blends, in which one of the SWs provides the blend with the primary stress and syllable number (*strujadin* ← st(ru)ja × St(o)jadin) (Ronneberger-Sibold 2006: 168). Although similar prosodic patterns may also occur in other blend types, contour blends are identified by the fact that this feature serves as the primary structuring principle of the blend. In line with Ronneberger-Sibold's (2006) typology, when multiple characteristics are present, classification is based on the feature that most strongly determines the formal structure of the blend. The (overlapping) contour blends in our sample are: *kvascinantno*, *strujadin*, *isvinjavam se*, *Medolada*, *Prasorog*, *Urmolada*, *žiraćino*, *Dveromobil*, *svinjoid*, *stresolovka*, *kešformisanje*.
5) (Overlapping) semi-complete blends (*Malimeta*, *rurbani*, *gurmetnica*, *Zvezdapol*, *lažovizija*, *šljamformer*, *zbunjoblak*, *zaštipci*, *kakaceza*, *ćelepinja*, *šazbuka*), in which one SW is full and the other one is clipped (*šazbuka* ← š(umska) × azbuka). Unlike the clipped word of contour blends, the clipped word of semi-complete blends does not provide the [blend's] "rhythmic contour" (Ronneberger-Sibold 2006: 169).
6) Fragment blends (*Gnjavnik*, *kornjačurke*, *Irkofas*, *reciklosaurusi*, *sočka*, *vevečka*, *žiramun*, *šumkula*, *ponštipci*), in which both SWs are reduced to fragments, which may overlap (Ronneberger-Sibold 2006: 169) (*vevečka* ← veve(rica) × (ma)čka).

5 Note that this typology is based on the perspective of blend creators, which does not necessarily correspond to that of readers or listeners.

6 All blends, along with their SWs and meanings, are listed in the Appendix.

The final dataset consists of 50 blends distributed across the six types as follows: 10 telescope, 5 inclusive, 4 superimposed, 11 contour, 11 semi-complete, and 9 fragment blends.

2.2 Questionnaire

The instrument used was a questionnaire. It was distributed in person by the author of this paper between June and November 2024. It was anonymous and divided into two parts, with the first part asking the informants to provide general information about themselves, including sex, age, education level, occupation, and type of residence area (urban vs. rural). Its second part included a list of 50 blends given in a decontextualized setting. The selection of blends is described in Section 2.1 above – the items were drawn from a larger corpus of 400 newly attested Serbian blends and chosen so as to include all six blend types, though not in equal proportions (see distribution in Section 2.1). Although the questionnaire did not provide a definition of the blending phenomenon, the process was exemplified by two blends – *zabare* and *lovčanik* (Bugarski 2019), together with their SWs. The informants were asked to identify the two SWs used to form each blend. The task was not time-limited, and it was completed on site under the author's supervision.

2.3. Informants

Our sample of informants comprised 272 native Serbian speakers from the Zlatibor District, Western Serbia, who participated in the research on a voluntary basis. Of the 272 informants, 104 were male and 168 were female. They were divided into four age groups: (a) 18–30 (59), (b) 31–44 (79), (c) 45–60 (83), (d) 61–75 (51).[7] The age groups were established post hoc, based on the distribution of participants across age ranges, in order to ensure relatively balanced group sizes and enable meaningful comparison across different stages of adulthood. Regarding the education level, the informants were also divided into four groups: (a) primary school (6), (b) three-year vocational school (18), (c) grammar school or four-year vocational school (125), (d) college or university (123). As many as 218 of the informants were urban residents, and 54 resided in rural areas. Regarding occupations, the available sample of 235 informants (13 questionnaires were incomplete, while 24 informants were retired or unemployed) was used for occupational analysis. Occupation was operationalized as the informants' current occupational status. Accordingly, the 24 informants who were retired or unemployed at the time of data collection were not included in the occupational category of active employment, as the study focused on current occupational status rather than full occupational

7 The number in brackets indicates the number of informants per group.

history. The remaining sample was divided into 10 occupational groups, relying on the typology provided by the Tertiary Collection of Student Information (TCSI) Field of Education Types: (a) natural and physical sciences (9), (b) information technology (4), (c) engineering and related technologies (60), (d) agriculture and environment (5), (e) health (15), (f) education (49), (g) management and commerce (35), (h) society and culture (35), (i) creative arts (10), and (j) food, hospitality and personal services (13). These categories served as the initial classification and were subsequently grouped into broader groups for statistical analysis (see Section 2.4), in order to ensure sufficient sample sizes and comparability across groups. For respondents who were still in secondary or tertiary education (high school pupils and university students), occupational categorization was based on their current educational program, which was used as an indicator of occupational group membership. The six respondents whose highest completed level of education was primary school were secondary-school students enrolled in an arts high school (as specified in the questionnaire) and were accordingly included in the 'creative arts' group.

2.4 Methods

Quantitative analyses, which involved calculating the percentages of correct answers for different blend types as well as for (non-)overlapping blends, were first performed in order to address RQ1 and RQ2, which investigate whether morphotactic transparency and overlapping as structural features facilitate source-word recognizability. A qualitative analysis of the questionnaire results was also conducted, focusing on specific examples of blends with the highest and lowest rates of successful reconstruction (operationalized as the proportion of correct responses per item and interpreted as an indicator of item difficulty), in order to illustrate patterns observed in the quantitative analysis and to provide additional insight into factors influencing reconstructability.[8] Both these analyses were conducted because we believe that together they can help us better understand why some blends lend themselves to more successful reconstruction than others.

The data, i.e., the correct answers for all 50 blends (regardless of their type), were also analyzed statistically in relation to the informants' sociodemographic variables (sex, age, education level, occupation, and type of residence area – rural vs. urban). These sociodemographic variables were included on the basis of previous sociolinguistic and psycholinguistic research (see, e.g., Halupka-Rešetar & Lalić-Krstin 2012; Lehrer 1996, 2003; Silaški & Đurović 2013), which suggests that lexical processing and word-formation awareness may vary across speaker groups with different educational and linguistic profiles and may therefore influence speakers' success in reconstructing the source words of blends. A statistical

8 Due to lack of space, most of them were not included in the paper.

analysis was performed using R version 4.2.0 (2022-04-22 ucrt) (R Core Team 2022). The level of significance was set to .05, thus *p*-values lower than .05 were considered statistically significant. The data in this study consist of count data of correct answers per item, where each item (i.e., each blend) constitutes one observation. A response was considered correct only if both source words corresponding to the blend were accurately identified. Following Halupka-Rešetar and Lalić-Krstin (2012), morphological variation (e.g., inflectional or derivational forms) was disregarded, provided that the underlying lexical bases were correctly identified and represented linguistically valid forms (i.e., not partial or non-independent elements). Observation no. 45 (*šazbuka*) was excluded from item-level analyses due to the absence of correct responses across all informants, which did not allow meaningful comparative analysis of item difficulty or reconstructability for this blend. Missing values, naturally occurring if no informant reconstructs the blend successfully, are imputed with zeros. Here, we are interested in testing whether the average amount of correct responses varies significantly among groups. Sample sizes among groups vary greatly and it is currently not possible to even out the data due to the limited reach of our study. For example, it is difficult to find informants from rural areas who have completed primary education only and are willing to complete the questionnaire. Therefore, to avoid misleading differences caused by unequal group sizes, we used percentages of correct answers instead of raw counts. Accordingly, we used the Kruskal-Wallis (KW) test (Kruskal and Wallis 1952) from the R package stats. The KW test is a non-parametric test suitable for skewed (non-normal) data, which is present in our case. We also calculated the effect size for the KW test. If the KW test rejects the null hypothesis that the average percentage of correct answers among groups is the same, we perform the post-factum analysis using the Dunn test (Dunn 1964) from the R package dunn.test to locate the exact difference among the groups. The Holm adjustment of *p*-value for multiple comparisons is also performed (Holm 1979).

To determine whether there are significant differences in the number of correct answers across occupational groups, we grouped the occupational categories defined in Subsection 2.3 into two broader groups, namely Humanities, Arts, and Social Sciences (HASS) (94 informants) and Science, Technology, Engineering, and Mathematics (STEM) (78 informants), due to the small and uneven sample sizes across individual subgroups and in order to obtain groups of more comparable size for this analysis. Additionally, we conducted a more fine-grained analysis within the larger occupational grouping by examining three subgroups with comparable sample sizes – Education (49), Management and Commerce (35), and Society, Culture and Arts (45) – in order to explore potential differences within the Humanities, Arts, and Social Sciences (HASS) domain. Consequently, the KW test was used to test for the difference in the average number of correct answers among these three groups. The

Наслеђе 63 • 2026 • 255–274

Mann-Whitney U-test (MW U-test) was implemented wherever the number of groups given was two. The effect measured in that case was the so-called rank-biserial correlation.

3 Results and discussion

3.1 Reconstruction of different blend types

The results of a quantitative analysis of correct answers show that the success in the reconstruction of different blend types is as follows: *telescope blends* (31%) → *superimposed blends* (21%) → *contour blends* (16%) → *semi-complete blends* (15%) → *fragment blends* (12%) → *inclusive blends* (5%). These findings support the view that the recognizability of source words is not categorical but gradient in nature, i.e., "a matter of degree" (Kemmer 2003: 77). The results are, in principle, compatible with the findings of Halupka-Rešetar and Lalić-Krstin (2012: 107) reported in Section 1, in that both studies show that blends with a higher degree of overlap and greater preservation of source-word material are more easily reconstructed, whereas those involving a substantial loss of material tend to be more difficult to decode, despite the fact that the two analyses are based on different typologies of blends. Considering that the percentage of the informants who correctly identified the SWs of inclusive blends is lowest, our first hypothesis is not completely confirmed. That is, although our results indicate that morphotactic transparency generally facilitates blend reconstruction, inclusive blends appear to be a major exception to this tendency. One possible explanation for this is that, in the absence of contextual cues, informants may fail to recognize that one source word is fully embedded within another, despite the graphic marking used to signal the included segment. As a result, the structural transparency of such blends may not translate into processing ease, since the recognition of inclusion requires an additional analytic step compared to other blend types. Unfortunately, it is not possible to compare our results for inclusive blends with those of Halupka-Rešetar and Lalić-Krstin (2012: 103) because they did not include examples of inclusive blends in their analysis. It is also noteworthy that Tomić (2021) found that there was a general tendency towards a more successful reconstruction of semi-complete and complete English blends (by non-native speakers) than of those produced by contour and fragment blending.[9] Although our research did not use a psycholinguistic approach to SW recognizability (as was the case in Lehrer's (1996; 2003) psycholinguistic experiments), it is worth noting that, contrary to our results, the SWs of English blends such as *cattitude* ← c<u>at</u> × <u>at</u>titude (a telescope blend in the terminology adopted for the present study) or *entreporneur* ←

9 It should be noted, however, that no distinction was made in that study between different subtypes of complete blends (e.g., telescope, inclusive, or superimposed blends), as proposed in the present research.

entrep(re)neur × porn(ography) (a contour blend in the terminology adopted for the present study), were least easily recognized among Lehrer's (2003: 371–372) data. It is also interesting to observe that, out of all telescope blends in our sample, *sarmada* (whose SWs share as many as four graphemes) was by far the most difficult to reconstruct. Specifically, most of the informants had difficulty recognizing its SW2 – *armada* ('armada'). One possible explanation for this may involve several interacting factors, including the relatively low frequency of the word *armada* in Serbian (AF=1,095; RF=0.38) (MaCoCu Serbian Web v1 (2021–2022)), the absence of contextual cues in the experimental setting, and the lack of a strong semantic association between *sarma* ('a Serbian traditional dish made of cabbage, minced meat, and rice') and *armada* ('armada').[10] Some of the most common incorrect answers for this SW2 were *pomada*, *nada*, *gromada*, and *mada*. What is also implied by these incorrect answers is that the greater number of lexical neighbors – i.e., words that share similar initial or final segments with the source word – is likely to negatively influence SW recognizability (see also Lehrer 1996; 2003; Silaški and Đurović 2013: 97).

Among overlapping semi-complete blends, *Malimeta*, in which SWs overlap continuously, was most difficult to reconstruct, mainly because the informants misidentified its SW1 – the noun *malina* 'raspberry' – for the adjective *mali* 'small' and its SW2 – *limeta* 'lime' – for the word *meta* 'target'. Regarding the most common incorrect answers, mention must also be made of the fact that *mali* (AF=437,545; RF=150.53) and *meta* (AF=61,163; RF=21.04) are considerably more frequent than *malina* (AF=18,851; RF=6.49) and *limeta* (AF=916; RF=0.32), respectively. Furthermore, these misidentifications may be partly accounted for by the *Principle of Least Effort* (Zipf 1949: 1), which states that people tend to solve problems "in such a way as to minimize the total work that [they] must expend in solving" them, and which in the present context suggests that speakers tend to favor more frequent and cognitively accessible forms. This is consistent with the fact that *mali* and *meta* are considerably more frequent than *malina* and *limeta*. At the same time, the absence of contextual cues and limited extralinguistic knowledge may also have contributed to the observed misidentifications. A similar tendency was observed in other examples (belonging to different blend types), such as the blend *Miškomor*, which was frequently decomposed into *miš* 'mouse' and *komora* 'chamber'.

The quantitative analysis of the correct answers for all blends in the sample, divided into overlapping and non-overlapping blends as two independent subsamples, shows that the SWs of 27 overlapping blends are more successfully identified than those of 23 non-overlapping blends. Overlapping blends

10 The abbreviations AF and RF stand for *absolute* and *relative frequency* (per million tokens), respectively. All frequencies are retrieved from the MaCoCu Serbian Web v1 (2021–2022), available in the Sketch Engine software (Sketch Engine). According to Kemmer (2003: 77), the factor of familiarity is correlated with the frequency factor.

include telescope, inclusive, and superimposed blends, in which overlap is a defining structural property. Contour, semi-complete, and fragment blends were not categorically treated as non-overlapping; rather, only those instances in which no overlap was present were included in the non-overlapping group, as overlap is not a defining structural property of these blend types, although it may occur in other instances. Specifically, 46% of the informants managed to correctly identify SWs of all overlapping blends in the sample, compared to 29% of those who succeeded in correctly identifying SWs of non-overlapping blends. This piece of evidence, which is also in line with Halupka-Rešetar and Lalić-Krstin's (2012) findings, allows us to verify our second hypothesis (for similar results as regards (non-)overlapping English blends, cf. Mattiello 2022 and Tomić 2021). These results are, however, contrary to the findings of Silaški and Đurović (2013: 97), who showed that, for non-native speakers, English "blends with partial or complete overlap are very difficult to process".

The results of the above linguistic analysis suggest that, besides the amount of material preserved from each SW and the presence of overlap (though not extensive), other factors such as SW frequency and the number of lexical neighbors may also play a role in facilitating blend reconstruction. However, since the present study did not systematically control for variables such as the number of lexical neighbors or the semantico-pragmatic relatedness of the source words, these factors should be interpreted cautiously. In particular, it may be hypothesized that blends whose source words belong to the same lexical field or are more closely related in the extralinguistic world (e.g., *malina* and *limeta*) are easier to process than those whose constituents are semantically less related (e.g., *sarma* and *armada*). Further psycholinguistic research, based on a larger and more controlled sample, is needed to test the contribution of these factors.

3.2 Influence of sociodemographic factors on blend reconstruction

A Kruskal–Wallis test revealed no statistically significant effect of age on the informants' accuracy in reconstructing the SWs of blends, $H(3) = 4.22$, $p = .24$, $\eta^2 = .022$, indicating a negligible effect. A Mann–Whitney U test revealed no statistically significant difference between male and female informants in their accuracy in reconstructing the SWs of blends, $U = 1135$, $p = .64$, $r = -.05$, indicating a negligible effect. A Kruskal–Wallis test revealed no statistically significant effect of education level on the informants' accuracy in reconstructing the SWs of blends, $H(3) = 2.786$, $p = .43$, $\eta^2 = .022$, indicating a negligible effect. A Mann–Whitney U test revealed no statistically significant difference between rural and urban informants in their accuracy in reconstructing the SWs of blends, $U = 1197$, $p = .98$, $r = -.003$, indicating a negligible effect. Although none of the *p*-values is statistically significant, the effect sizes suggest relatively stronger associations for *age* and *education level* than for *sex* and *area of residence*, which

may indicate a slightly greater influence of these variables on blend reconstruction (cf. Halupka-Rešetar & Lalić-Krstin 2012).

As far as the *occupation* variable is concerned, the results of the statistical analysis for two groups (HASS and STEM) are as follows: A Mann–Whitney U test revealed no statistically significant difference between the groups, $U = 991.5$, $p = .07$, $r = -.21$, indicating a small to medium effect. In this case, it was appropriate to compare raw scores, as the two groups did not differ substantially in their distributions or show marked imbalance. Therefore, there was no indication that these differences would affect the comparison of group performance. This contrasts with the previous analyses involving age, sex, education level, and area of residence, where strongly uneven group sizes required the use of rank-based non-parametric tests.

However, in a finer-grained analysis of the possible influence of occupation on blend reconstruction, the original 10 occupational categories were combined into three broader groups (Education (E) (49), Management and Commerce (MC) (35), and Society, Culture and Arts (SCA) (45)) in order to ensure sufficiently large and statistically comparable group sizes. A Kruskal–Wallis test revealed a statistically significant difference among the three groups, $H(2) = 7.936$, $p = .02$, indicating a meaningful effect of occupation on blend reconstruction. Although the result is statistically significant, $\eta^2 = .054$ indicates a small effect. To investigate which specific groups differed significantly from one another, Dunn's post-hoc test was conducted following the KW test. The Dunn post-hoc test (Table 1) revealed a significant difference between the Education and Management and Commerce groups (adjusted $p = .027$), with the Vargha–Delaney A effect size (A = .646) (Vargha and Delaney 2000) indicating a medium effect and higher reconstruction accuracy in the Education group.

Table 1. The results of Dunn's post-hoc test and Vargha-Delaney A effect size measures

Comparison	Z	Unadjusted *p*-value	Adjusted *p*-value	Effect size	Interpretation
E vs. MC	2.611	.009	.027	.646	Medium effect
E vs. SCA	0.391	0.696	.696	.529	Small effect
MC vs. SCA	-2.221	0.026	.053	.364	Small effect

It can be inferred from the above results that the only statistically significant difference in occupational performance was observed between the Education and Management and Commerce groups, with higher reconstruction accuracy in the Education group. No statistically significant differences were found between the Education and Society, Culture and Arts groups (adjusted $p = .696$), nor between the Management and Commerce and Society, Culture and Arts groups after adjustment, indicating that these groups performed at a broadly comparable

Наслеђе 63 • 2026 • 255–274

level. While the Education group shows a tendency toward higher performance, this pattern should be interpreted with caution, as most pairwise comparisons do not reach statistical significance. A tentative explanation in terms of occupational creativity may be considered, but it cannot be firmly established on the basis of the present data. Overall, these results do not allow firm conclusions regarding systematic differences across all occupational groups, but suggest that the occupational domain may play a role in blend reconstruction.

Finally, the results of the statistical analysis provide partial support for our third hypothesis. As expected, no statistically significant effects were found for sex, age, education level, or area of residence, suggesting that these sociodemographic factors do not have a significant influence on blend reconstruction. By contrast, the analysis of occupational groups revealed one statistically significant difference between the Education and Management and Commerce groups, which is in line with the expectation that occupational variation may play a role, possibly due to differences in linguistic creativity and exposure to language use in specific professional contexts.

4 Conclusion

The quantitative analysis of the informants' success at reconstructing the SWs of Serbian blends produced by six different blending techniques shows that success in blend reconstruction, in general, positively correlates with the relative morphotactic transparency of blends, which is in line with the expectation that more transparent morphological structures are easier to process, except in the case of inclusive blends which, despite being the second most transparent type, proved to be most difficult to reconstruct, possibly because they rely primarily on orthographic rather than phonological or semantic cues. General tendencies emerging from the quantitative analysis of overlapping and non-overlapping blends are that the former type is more successfully reconstructed than the latter, which is in line with the expectation that overlapping structures provide more segmental cues and thus facilitate the identification of source words. The qualitative analysis of answers for individual examples indicates that, in addition to the aforementioned factors, some other elements (e.g., greater frequency of SWs or the number of their lexical neighbors) may also facilitate blend reconstruction. Based on these results, we can conclude that blend reconstruction is a complex matter and also agree with Kemmer (2003: 77) that "recognizability is a matter of degree". The statistical analysis of the influence of a number of sociodemographic factors on the informants' success in blend reconstruction shows that there are no statistically significant differences in the percentages of correct answers among the investigated sex, age, education, and urban or rural residence groups. However, the investigation of the influence of different occupational groups on blend reconstruction

shows that there is a significant difference between the Education group and the Management and Commerce group, which may be explained by differences in linguistic creativity, as well as exposure to language use in these occupational settings.

Considering that products of lexical blending in Serbian still seem to be a puzzle for its native speakers, many of whom approach the reconstruction of blends as if they were compounds, further research could focus on conducting a longitudinal study into the recognizability of Serbian blends' SWs in order to track possible changes in the tendencies observed herein. It would also be interesting to compare the results of the statistical analysis with the answers of informants from other districts. Although one might not expect substantial differences, given the increasing similarity in linguistic exposure through media and the Internet, such a comparison could nevertheless reveal more subtle regional or sociolinguistic variation in blend processing. Taking into account the occupation-related results and the fact that blends are highly creative formations, additional analyses could explore the influence of other occupations on blend reconstruction or the role of creativity in SW recognizability, using creative thinking tests.

SOURCES

Tomić 2023: G. Tomić, *Leksičke slivenice u savremenom engleskom i srpskom jeziku: strukturni, sadržinski i sociolingvistički aspekti* [Lexical Blends in Contemporary English and Serbian: Structural, Semantic, and Sociolinguistic Aspects]. Neobjavljena doktorska disertacija [Unpublished PhD dissertation]. Kragujevac: Filološko-umetnički fakultet.

REFERENCES

Adams 1973: V. Adams, *An Introduction to Modern English Word-Formation*, London/New York: Longman.

Bauer 2012: L. Bauer, Blends: Core and periphery, *in*: V. Renner, F. Maniez, and P. J. Arnaud (eds.), *Cross-Disciplinary Perspectives on Lexical Blending*, Berlin/New York: De Gruyter Mouton, 11–22. https://doi.org/10.1515/9783110289572.11

Bauer 2017: L. Bauer, *Compounds and Compounding*, Cambridge: New York: Cambridge University Press.

Beliaeva 2019: N. Beliaeva, Blending in Morphology, *in:* M. Aronoff (ed.), *Oxford Research Encyclopedia of Linguistics*, Oxford: Oxford University Press, n.p. https://doi.org/10.1093/acrefore/9780199384655.013.511

Bugarski 2001: R. Bugarski, Dve reči u jednoj: leksičke skrivalice [Two Words in One: Lexical Puzzles], *Jezik danas*, 5(13), 1–5.

Bugarski 2016: R. Bugarski, Slivenice kao pokazatelj promena u jeziku i društvu [Blends as Indicators of Changes in Language and Society], *u*: J. Dražić, I. Bjelaković i D. Sredojević (ur.), *Zbornik u čast Ljiljani Subotić: Teme jezikoslovne u srbistici kroz dijahroniju i sinhroniju*, Novi Sad: Filozofski fakultet, 511–529.

Bugarski 2019: R. Bugarski, *Srpske slivenice: monografija sa rečnikom* [Serbian Blends: A Monograph with a Dictionary], Novi Sad: Akademska knjiga.

Dressler 2000: W. U. Dressler, Extragrammatical vs. Marginal Morphology, *in:* U. Doleschal and A. M. Thornton (eds.), *Extragrammatical and Marginal Morphology*, München: Lincom Europa, 1–10.

Dressler 2005: W. U. Dressler, Word-Formation in Natural Morphology, *in*: P. Štekauer and R. Lieber, *Handbook of Word-Formation*, Dordrecht: Springer, 267–284.

Dunn 1964: O.J. Dunn, Multiple comparisons using rank sums, *Technometrics*, 6, 241–252.

Fandrych 2008: I. Fandrych, Submorphemic Elements in the Formation of Acronyms, Blends and Clippings, *Lexis: Journal in English Lexicology*, 2, 105–122. https://doi.org/10.4000/lexis.713

Halupka-Rešetar & Lalić-Krstin 2012: S. Halupka-Rešetar & G. Lalić-Krstin, Razumevanje slivenica u srpskom jeziku [Interpreting blends in Serbian], *Nasleđe*, Kragujevac, 9 (22), 101–109.

Halupka-Rešetar and Lalić-Krstin 2009: S. Halupka-Rešetar and G. Lalić-Krstin, New blends in Serbian: typological and headedness-related issues, *Godišnjak Filozofskog fakulteta u Novom Sadu*, XXXIV-1, 115–124.

Holm 1979: S. Holm, A simple sequentially rejective multiple test procedure, *The Scandinavian Journal of Statistics*, 6(2), 65–70.

Kemmer 2003: S. Kemmer, Schemas and lexical blends, *in:* H. Cuyckens et al. (eds.), *Motivation in Language: From Case Grammar to Cognitive Linguistics, Studies in Honour of Günter Radden, Amsterdam/Philadelphia: John Benjamins*, 69–97. https://doi.org/10.1075/cilt.243.08kem

Kruskal and Wallis 1952: W. H. Kruskal and W. A., Use of ranks in one-criterion variance analysis, *Journal of the American Statistical Association*, 47.260, 583–621.

Lalić-Krstin & Halupka-Rešetar 2007: G. Lalić-Krstin & S. Halupka-Rešetar, Nešto novo o novim slivenicama u srpskom jeziku [Some New Insights into New Blends in the Serbian Language], *Svet reči: srednjoškolski časopis za srpski jezik i književnost*, 11(23/24), 26–30.

Lehrer 1996: A. Lehrer, Identifying and interpreting blends: An experimental approach, *Cognitive Linguistics*, 7(4), 359–390. https://doi.org/10.1515/cogl.1996.7.4.359

Lehrer 2003: A. Lehrer, Understanding trendy neologisms, *Italian Journal of Linguistics*, 15(2), 369–382.

Mattiello 2013: E. Mattiello, *Extra-Grammatical Morphology in English: Abbreviations, Blends, Reduplicatives, and Related Phenomena*, Berlin/Boston: De Gruyter Mouton.

Mattiello 2022: E. Mattiello, Blend recognizability in English as a foreign language: An experiment, *Lingue Linguaggi*, 54, 31–51.

R Core Team (2022). *R: A language and environment for statistical computing*. R Foundation for Statistical Computing, Vienna, Austria. https://www.R-project.org/.

Renner 2015: V. Renner, Lexical Blending as Wordplay, *in*: A. Zirker and E. Winter Froemel (eds.), *Wordplay and Metalinguistic/Metadiscursive Reflection: Authors, Contexts, Techniques, and Meta-Reflection*. Berlin: De Gruyter Mouton, 119–134. https://doi.org/10.1515/9783110406719-006

Renner 2023: V. Renner, Blending, *in*: Ackema, Peter & Bendjaballah, Sabrina & Bonet, Eulàlia & Fábregas, Antonio (eds.), *Wiley Blackwell Companion to Morphology*, Wiley-Blackwell: Hoboken, 251–270. https://doi.org/10.1002/9781119693604.morphcom009.

Ronneberger-Sibold 2006: E. Ronneberger-Sibold, Lexical Blends: Functionally Tuning the Transparency of Complex Words, *Folia Linguistica*, XL, 40(1–2), 155–181.

Silaški and Đurović 2013: N. Silaški and T. Đurović, Of 'Siliconaires' and 'Millionerds' – How ESP learners understand novel blends in English, *Ibérica*, 25, 85–106.

Sketch Engine. www.sketchengine.eu/.

TCSI. Field of Education Types. TCSI Support. Available at www.tcsisupport.gov.au/resources/field-of-education-types. Accessed 18 Dec. 2024.

Tomić 2019: G. Tomić, Tvorbeno-semantička analiza novih slivenica u srpskom jeziku [A Morphosemantic analysis of new blends in Serbian], *Zbornik radova Filozofskog fakulteta u Prištini*, XLIX (2), 61–83. https://doi.org/10.5937/zrffp49-21405

Tomić 2021: G. Tomić, Tumačenje slivenica u nazivima brendiranih proizvoda jela i pića na engleskom jeziku [Interpreting blends in food and drink brand names in English], *u*: M. Kovačević i J. Petković (ur.), *Savremena proučavanja jezika i književnosti: zbornik radova sa XII naučnog skupa mladih filologa Srbije*, Knj. 1, Kragujevac: Filološko-umetnički fakultet, 79–97.

Tomić 2022: G. Tomić, Lexical Blends in Serbian: An Analysis of Morphosemantic Transparency. *Zbornik radova Filozofskog fakulteta u Prištini*, LII (3), 13–37. https://doi.org/10.5937/zrffp52-32090

Tomić 2023: G. Tomić, *Leksičke slivenice u savremenom engleskom i srpskom jeziku: strukturni, sadržinski i sociolingvistički aspekti* [Lexical Blends in Contemporary English and Serbian: Structural, Semantic, and Sociolinguistic Aspects]. Neobjavljena doktorska disertacija [Unpublished PhD dissertation]. Kragujevac: Filološko-umetnički fakultet.

Vargha and Delaney 2000: A. Vargha and H. D. Delaney, A Critique and Improvement of the CL Common Language Effect Size Statistics of McGraw and Wong, *Journal of Educational and Behavioral Statistics*, 25(2), 101–132. https://doi.org/10.3102/10769986025002101

Zipf 1949: G.K. Zipf, *Human Behavior and the Principle of Least Effort*, Cambridge: Addison-Wesley Press.

ACKNOWLEDGMENTS

I thank the two anonymous reviewers for their invaluable comments on an earlier version of the manuscript.

Наслеђе 63 • 2026 • 255–274

Горица Р. Томић

РЕКОНСТРУИСАЊЕ НОВИХ ЛЕКСИЧКИХ СЛИВЕНИЦА У СРПСКОМ ЈЕЗИКУ: (СОЦИО)ЛИНГВИСТИЧКИ НАЛАЗИ

Резиме

Циљ овог рада је да анализира успешност изворних говорника српског језика у реконструкцији шест различитих типова српских сливеница, као и то да ли преклапање, као дистинктивно обележје творбеног поступка сливања, доприноси успешнијој реконструкцији сливеница. Рад такође има за циљ да истражи да ли постоје статистички значајни утицаји неких социодемографских фактора на успешност у реконструисању сливеница. Резултати квантитативне анализе показују да испитаници најуспешније реконструишу телескопске, а најмање успешно инклузивне сливенице, као и да преклапање, у начелу, доприноси успешнијој реконструкцији. Анализа утицаја социодемографских фактора (године живота, пол, образовање, занимање и тип средине – урбана/рурална) показала је да не постоји статистички значајан утицај фактора пола, година живота, нивоа образовања и средине на успешност у препознавању елемената сливенице, али да постоје статистички значајне разлике у броју тачних одговора између одређених група занимања, што се може објаснити различитим степеном креативности који карактерише ова занимања.

Кључне речи: *лексичке сливенице, српски језик, изворни говорници, реконструкција сливеница*

Примљен: 1. марта 2025. године
Прихваћен: 30. априла 2026. године

Appendix

Blend	SW1		SW2		Possible blend interpretation
1. dnovinarka	dno	'bottom, lowest point'	novinarka	'female journalist'	'a low-quality female journalist'
2. tviteroristkinja	Tviter	'Twitter'	teroristkinja	'female terrorist'	'a woman who aggressively attacks others on Twitter'
3. hranac	hrana	'food'	ranac	'rucksack'	'a rucksack for carrying food'
4. trološ	trol	'troll (the Internet)'	ološ	'scum'	'a troll who is also scum; an especially nasty troll'
5. tviterapeut	Tviter	'Twitter'	terapeut	'therapist'	'someone who acts as a therapist on Twitter'
6. zloslobođenje	zlo	'evil'	oslobođenje	'liberation'	'a 'liberation' that results in harm'
7. kondukterijer	kondukter	'bus conductor'	terijer	'terrier (dog)'	'creatures that have the memory of a ticket inspector and the persistence and energy of a terrier'
8. muzikaljenje	muzika	'music'	kaljenje	'forging'	'the process of being shaped through music'
9. sarmada	sarma	'stuffed cabbage rolls'	armada	'army'	'an invincible armed force whose power is based on the superior energy of its soldiers, brought about by their regular consumption of the delicious dish made of cabbage, minced meat, and rice'
10. grokenrol	grok	'oink'	rokenrol	'rock and roll'	'rock and roll performed in a pig-like, grunting manner'
11. doČEPaj se	dočepaj se	'grab/get hold of'	čep	'cork'	'grab a cork (for donation)
12. NINdže	NIN	'NIN (magazine)'	nindže	'ninjas'	'journalists of NIN seen as 'ninjas''
13. nazaDNO	nazadno	'regressive'	dno	'bottom'	'backward to the point of being bottom-level'
14. BORba	BOR	'BOR (company name)'	borba	'struggle'	'a struggle associated with BOR'

15. GIMnastičar	GIM	'GIM (company name)'	gimnastičar	'gymnast'	'A minister perceived as a 'gymnast' (figuratively), due to alleged irregularities involving the arms-producing company GIM'
16. Pivonirka	pivo	'beer'	pionirka	'female pioneer'	'name of a beer'
17. Miškomor	Miško	'male name'	mišomor	'rat poison'	'something harmful associated with Miško'
18. zlaova	zla	'evil'	zaova	'sister-in-law'	'an evil sister-in-law'
19. Svetlozar	svetlo	'light'	Svetozar	'male name'	'name of a glowworm character in a cartoon'
20. kvascinantno	kvas	'kvass'	(f)ascinantno	'fascinating'	'fascinating in a kvass-related way'
21. strujadin	st(ru)ja	'electricity'	St(o)jadin	'car model'	'an electric version of a Stojadin (car) (humorous)'
22. isvinjavam se	i(z)vinjavam se	'I apologize'	svinja	'pig'	'an apology expressed in a 'pig-like' manner; I pigologize'
23. Malimeta	mali(na)	'raspberry'	limeta	'lime'	'name of a drink; raspberry–lime beverage/flavor blend'
24. rurbani	rura(l)ni	'rural'	urbani	'urban'	'semi-rural, semi-urban'
25. gurmetnica	gurm(an)	'gourmet'	umetnica	'female artist'	'a culinary artist; gourmet artist (female)'
26. Gnjavnik	gnjav(i)	'to annoy'	(dne)vnik	'news'	'annoying news'
27. kornjačurke	kornjač(e)	'turtles'	(pe)čurke	'mushrooms'	'turtle-shaped mushrooms / hybrid creatures'
28. Medolada	med	'honey'	(č)okolada	'chocolate'	'a unique cream with honey, hazelnuts, and cocoa powder'
29. Prasorog	pras(e)	'piglet'	(jedn)orog	'unicorn'	'a pig-unicorn hybrid'
30. Urmolada	urm(e)	'dates (fruit)'	(č)okolada	'chocolate'	'a spread that represents a combination of dates with other fruits'
31. žiraćino	žir	'acorn'	(k)apućino	'cappuccino'	'an acorn-based cappuccino-like drink'
32. Dveromobil	Dver(i)	'political movement'	(aut)omobil	'car'	'a vehicle associated with Dveri'

33. svinjoid	svinj(a)	'pig'	(tabl)oid	'tabloid'	'a 'pig-like' tabloid'
34. stresolovka	stres	'stress'	(miš)olovka	'mousetrap'	'a stress trap; something that relieves stress'
35. Irkofas	Irko(m)	'company name'	fas(ada)	'facade'	'Facade paint produced by Irkom'
36. reciklosaurusi	recikl(aža)	'recycling'	(din)osaurusi	'dinosaurs'	'recycling dinosaurs – It's a creative, child-friendly term used to name characters who teach kids about sorting waste and protecting the environment'
37. Zvezdapol	Zvezda	'company name'	pol(udisperzija)	'dispersion paint'	'a dispersion paint produced by Zvezda'
38. lažovizija	laž -o-	'lie'	(tele)vizija	'television'	'television that spreads lies'
39. šljamformer	šljam	'scum'	(In)former	'name of a tabloid newspaper'	'Informer portrayed as scum media'
40. zbunjoblak	zbunj(en)	'confused'	oblak	'cloud'	'a confused cloud'
41. zaštipci	za	'for'	(u)štipci	'fritters'	'fritters for someone'
42. kakaceza	kaka	'poop'	(prin)ceza	'princess'	'a princess who poops'
43. kešformisanje	keš	'cash'	(in)formisanje	'informing'	'A database and project by the Raskrikavanje analyzing media co-financing in Serbia'
44. ćelepinja	će(vap)	'kebab'	lepinja	'flatbread'	'flatbread with ćevapi'
45. šazbuka	š(umska)	'forest'	azbuka	'alphabet'	'a forest alphabet'
46. sočka	so(va)	'owl'	(ma)čka	'cat'	'an owl-cat hybrid'
47. vevečka	veve(rica)	'squirrel'	(m)ačka	'cat'	'a squirrel-cat hybrid'
48. žiramun	žira(fa)	'giraffe'	(maj)mun	'monkey'	'a giraffe-monkey hybrid'
49. šumkula	šum(ska)	'forest'	(aj)kula	'shark'	'a forest shark'
50. ponštipci	pon(eti)	'to take along'	(u)štipci	'fritters'	'take-away fritters'