

Chapter 15

Dual preservation in Slovenian: The verb supports the noun in semi-spontaneous production

Matic Pavlič^a, Artur Stepanov^b & Penka Stateva^b

^aFaculty of Education – University of Ljubljana ^bCenter for Cognitive Science of Language – University of Nova Gorica

We investigated the use of dual morphology on nouns in six Slovenian dialects which either completely or only partially preserved dual marking on the verb in a language-production experiment using a picture description task with a pre-set vocabulary. We compared our results against the typology of Slovenian dialects with respect to preserving the dual feature, as presented in Jakop (2008), and based on data from the Slovenian Linguistic Atlas (collected 1946–1999). We found that the use of the targeted dual of a noun in Slovenian dialects is influenced by the dual of a verb via agreement. More specifically, the higher rate of preserved verb dual forms are associated with higher use of the dual and lower use of the plural in the subject – but not in the object. Greater use of the dual in the subject does not affect greater use of the dual in the object – although the nominative and accusative forms are identical in the masculine. These findings, for the first time, experimentally confirmed the previously suspected (Tesnière 1925) supporting role of verb agreement on preservation of the dual in Slovenian.

1 Introduction

This paper presents an experiment that seeks to identify and document the diachronic process of the loss of dual morphology in selected Slovenian dialects against Jakop's (2008) typological observations that are based on the Slovenian Linguistic Atlas (collected 1946–1999). To this end, we compared the proportion



of false plural forms instead of target dual ones, on the nouns as well as the verb in unmarked SVO sentences. We further ask whether the loss of dual morphology in nouns correlates with their grammatical function or, in other words, with the structural position of subject vs. object. To this end, we also compared the proportion of false plural forms instead of target dual as a function of the structural position of the noun in question. Our experimental results confirmed the earlier assumption that the dialects in which the dual in verbs is more preserved also preserve more dual morphology in (subject) nouns than the dialects in which the dual in verbs is less preserved. The subject-object asymmetry in dual preservation is attributed to the fact that subject-verb agreement facilitates the preservation of dual morphology in subject nouns.

1.1 Grammatical number

The vast majority of languages that encode grammatical number on a noun distinguish between singular and plural (e.g. English in Table 1), but there are also languages with a three-part distinction (e.g. Slovenian and some other Indo-European, Semitic, Austronesian, and South American languages). On the other hand, languages with a four- or five-part distinction are extremely rare.

Table 1: Number on English, Arabic and Slovenian noun for ‘student’.

	English	Arabic	Slovenian
Singular	student-Ø	taalib-Ø	študent-Ø
Dual	/	taalib-een	študent-a
Plural	student-s	taalib-iin	študent-i

The dual is marked compared to the plural with respect to several criteria: children acquire the dual later than the plural (Ravid & Hayek 2003), having the dual in a language entails having the plural (Greenberg 1966),¹ and the use of the dual tends to decrease in diachronic change (Corbett 2000). We can also add the morphological criterion of markedness, as mentioned in Greenberg’s Universal number 35 (1966: 94):

There is no language in which the plural does not have some nonzero allomorphs, whereas there are languages in which the singular is expressed only by zero. The dual and the trial are almost never expressed only by zero.

¹“No language has a trial number unless it has a dual. No language has a dual unless it has a plural.” (Greenberg 1966: 94).

The markedness of a number can trigger changes that lead to neutralization, i.e. unification of the marked grammatical number with the less marked one (see, e.g. the *Markedness-triggered impoverishment hypothesis*, Nevins 2011; the *Morphosyntactic Feature Economy hypothesis*, Slobodchikoff 2019; or the *Diachronic model of the loss of dual in the context of minimalist syntax*, Stepanov & Stateva 2018). Moreover, we know from research on language acquisition under suboptimal conditions (i.e., through quantitatively and/or qualitatively limited contact with the first language) that grammatical number (and agreement processes in general) is particularly volatile or susceptible to change (Polinsky 2018), e.g. paradigm simplification (Berdicevskis & Semenuks 2022). However, a factor in the gradual loss of dual forms could also be the interaction or loss of other grammatical categories, e.g. the neutralization of certain inflectional forms within a single grammatical number, which would intuitively lead speakers to retain only the grammatical number that differentiates between multiple inflectional forms (Ivanov 1983). Tesnière (1925), on the basis of diachronic data from Indo-European languages, made generalizations regarding the order of pluralization of dual forms, according to: **inflection** (locative > genitive > dative > nominative/accusative), **gender** (feminine > neuter > masculine), **part of speech** (adjective > demonstrative > noun > numeral > personal pronoun), **grammatical function** (object > subject).²

1.2 Dual in Slovenian

In languages with overtly expressed grammatical number, the latter may be encoded by a free or bound morpheme, a modification of the root, and/or by a substitute root. Finally, it may also be phonologically unexpressed – and in the latter case identifiable only via secondary marking, namely, via matched features on the dependent items because of agreement. In inflectional languages such as Slovenian, grammatical number is part of a formal system of agreement involving an agreement *target* and an agreement *controller* (Corbett 2000) either within a noun phrase (Toporišič 2000: 109) or within a tense/inflectional phrase (Toporišič 2000: 608). Thus, speakers of Slovenian determine the form of the demonstrative ('these'), numeral ('two') and adjective ('old') in relation to the form of the noun ('lorries') in a noun phrase like (1). Similarly, they determine the verbal form ('overtake') in relation to the form of the noun phrase ('lorry' or 'campers'), which serves as the subject in a sentence like (2a) or (2b).

²">" stands for "is followed by" on the diachronic trajectory.

- (1) Ta dva stara tovornjaka
these.DU two.DU old.DU lorries.DU
'these two old lorries'
- (2) a. Prikolice prehiteva TOVORNJAK.
campers.F.PL.ACC overtake.SG lorry.M.SG.NOM
'It is the campers that the lorry overtakes.'
- b. Prikolice prehitevajo tovornjak.
campers.F.PL.NOM overtake.PL lorry.M.SG.ACC
'Campers overtake the lorry.'

Here, agreement serves as a crucial clue to determine the sentence's meaning, since both nouns are potential candidates for the agent thematic role. The ambiguity is rooted in the flexible word order in Slovenian and in the homophonous forms of nominative and accusative both in masculine singular and feminine plural. Thus, in example (2a), the addressee recognizes the singular noun 'lorry' as the subject on the basis of singular form of a verb, while in example (2b), the plural noun 'camper' is recognized as the subject – again based on the verbal features. As can be seen from examples (2a) and (2b), in Slovenian the grammatical categories of gender and number are encoded in a single ending, but there is both theoretical and experimental evidence that these are distinct features. In Slovenian, grammatical number must be expressed on the noun phrase, be it a noun (3a) or a personal pronoun (3b). This being said, the grammatical number is not morphologically expressed on the coordinated noun phrases involving proper names (3c) and, trivially, on the silent *pro* (3d); note that Slovenian is a *pro*-drop language. In the latter two examples, the number is reflected in the verbal features as an instance of agreement.

- (3) a. Otroka špricata teto.
kids.DU spray.DU aunt.SG
'The kids spray an/the aunt.'
- b. Onadva špricata teto.
they.DU spray.DU aunt.SG
'They spray an/the aunt.'
- c. Jan in Rok špricata teto.
Jan and Rok spray.DU aunt.SG
'Jan and Rok spray an/the aunt.'
- d. *pro* špricata teto.
spray.DU aunt.SG
'They spray an/the aunt.'

In the Slovenian dialects as well as in the standard variety, agreement between the verb and the subject is obligatory (Toporišič 2000: 271): if the subject is marked for plural, the dual form of the verb is ungrammatical (4a); if the subject is marked for dual, the plural form of the verb is ungrammatical (4b).

- (4) a. {Otroci / oni / Jan, Rok in Bor} {špricajo / *špricata} teto.
 kid.PL they.PL Jan Rok and Bor spray.PL spray.DU aunt.SG
 ‘{The kids / they / Jan, Rok and Bor} spray a/the aunt.’
- b. {Otroka / onadva / Rok in Bor} {špricata / *špricajo} teto.
 kid.DU they.DU Rok and Bor spray.DU spray.PL aunt.SG
 ‘{The two kids / they two / Rok and Bor} spray a/the aunt.’

1.3 Preservation of dual in Slovenian dialects

In examining the preservation of dual forms in Slovenian dialects, we follow the results of Jakop (2008), based on the material of the Slovenian Linguistic Atlas (SLA, Benedik 1999: 15). The questionnaire of SLA consisted of 870 numbered questions with sub-questions eliciting up to 2000 linguistic expressions from each informant. Mostly, the questions were expressions (nouns and verbs) in standard Slovenian, and the informants’ task was to translate them into their dialect and, in some cases, to give the full paradigm of the expression (Benedik 1999: 15). Some expressions were elicited by picture naming or by asking specific questions about the target (“What is it ...?”).

The material was collected between 1946 and 1999 in 413 locations within the borders of present-day Slovenia and in neighboring countries with Slovenian minorities. Each location was represented by three informants (a man, a woman, and a child under 14). It was reported that speakers with the most authentic dialect language (children) and speakers with the most developed metalinguistic understanding (teachers) dominate, but more than a quarter of the total material contains no information about the informants (Kenda-Jež 2002: 154). The main shortcoming of the material is its non-homogeneity, which is due not only to the long years of collecting and the different collectors, but also to the different qualifications of the collectors, the imprecision of the questions, the methodology and the transcription. Therefore, from several hundred questions, Jakop (2008) selected only 10 questions useful for the study of the dual. Note, that answers to these questions are not available for all the 413 locations. Below, we present the results by giving the percentage of retained dual forms for three exemplar nouns (one per gender).

The dual noun forms in Slovenian dialects are most often preserved in the masculine gender (96%; *brat* ‘brother’, $N_{loc}=275$; Jakop 2008: 135).³ Only half of the dialects preserve feminine dual forms (51%; *krava* ‘cow’, $N_{loc}=324$; Jakop 2008: 135–136). Note that neuter nouns are masculinized (and sometimes feminized) in 41% of the dialects (*okno* ‘window’, $N_{loc}=237$; Jakop 2008: 136), so that the neuter dual forms are retained only in 10% (*okno* ‘window’, $N=237$; Jakop 2008: 136). In addition, both masculine and feminine dual verb forms have been partially lost in Northern and Southern dialectal groups which leads Jakop (2008) to link dual loss to contacts with neighboring languages without dual (Italian and Croatian).⁴ In these dialects speakers use plural forms instead of the dual to describe an event with a participant consisting of two entities (see Figure 1 left) – as in our hypothetical example (5a). According to our informal observations this is especially common with a coordinated noun phrase (5b), a third person plural pronoun (5c), or the silent personal pronoun *pro* (5d) as a subject.

- (5) a. Otroci špricajo teto.
 kid.3.PL sprej.3.PL aunt.SG
 ‘The kids sprej a/the aunt.’
 b. Jan in Rok špricajo teto.
 Jan and Rok sprej.3.PL aunt.SG
 ‘Jan and Rok sprej a/the aunt.’
 c. Oni špricajo teto.
 they.3.PL sprej.3.PL aunt.SG
 ‘They sprej a/the aunt.’
 d. *pro*_i špricajo_i teto.
 sprej.3.PL aunt.SG
 ‘They sprej a/the aunt.’

Finally, we should also note that the earliest evidence of Slovenian losing dual can be found in the first printed books (16th century), as reported by Derganc (2006), Jakop (2008) and Orel (2019), a.o. However, to date “*the geographical prevalence of the use of dual forms in Slovenian dialects has not decreased significantly – the dual is a productive and living category in Slovenian.*” (Jakop 2008: 145).

³The number given by N_{loc} represents the number of locations for which a response to a SLA question was available (out of 413 locations).

⁴It is not clear, though, why contacts with German on the North does not lead to the dual loss.

2 Experiment

Our goals were to (i) verify the pattern of usage of the dual marking on nouns relative to the (in)complete loss/retention of verbal dual marking in Slovenian dialects (as reported in Jakop's 2008 work) in a controlled production experiment using a larger and more diversified set of uniform materials and a unified procedure; (ii) document potential dynamics of any existing or ongoing loss of the dual morphology in these dialects or a subset thereof; and (iii) explore the distribution and potential loss of the dual morphology relative to the noun's grammatical function. The rationale for the third goal was that most of the previous studies of the loss of dual have focused on nouns and verbs largely independently. However, the existence of morphological manifestation of subject-verb agreement, but not object-verb agreement, in Slovenian suggests that if the loss of the dual is taking place, it may be affecting subjects and objects to a different extent because of the asymmetry regarding their respective association with the verb. If verbs retain the dual form this might reinforce the dual on the subject longer than on the object, because of this association, as has been previously noticed by Tesnière (1925). We were therefore interested in whether this asymmetry has a systematic character that could be detected in a production study.

We used a picture description task and further restricted the informants' utterances by furnishing them with three key words from which they had to form a transitive sentence. This way we were able to check whether their use of dual nouns was related, on the one hand, to the extent of preservation of the nominal morphology in Slovenian dialects according to the data of SLA and the analysis of Jakop (2008) and, on the other hand, to the noun's sentential function as subject or object. We decided to study only the masculine nouns as those are considered the most stable group of nouns in Slovenian dialects (see above) and in the Indo-European languages in general (Tesnière 1925). For the same reasons, we focused on the syntactic functions of the subject and object, which are in transitive sentences encoded by nominative and accusative case, respectively.

Since Slovenian dialects (with the possible exception of southwestern dialects and southern dialects) do not allow masculine noun phrase in subject position not to agree with a verb in number (Jakop 2008), we hypothesize, in line with previous research (Tesnière 1925), that subject-verb agreement supports the use of dual and possibly contributes to the preservation of dual morphology in nouns. In our experiment, a more preserved verbal dual morphology in a dialect would result in more dual forms in the subject compared to the object – while a less or non-preserved verbal dual morphology in a dialect would result in a balanced use of the dual in the subject and object.

2.1 Dialects and informants

Table 1 shows the dialects we selected for our research: according to Jakop (2008) all of them have preserved dual noun morphology in nominative/accusative masculine (A and B dialects) while the dual verb morphology is lost in Southern and Western dialects (B dialects) but retained in Pohorje mountains, Soča river and Upper Carniola dialects (A dialects). In selecting the dialects, we used Jakop's (2008) two maps, based on two lemmas from SLA: a regular noun (*brat* 'brother') and a regular verb (*delati* 'to work'). We selected only regular expressions, since irregular paradigms are often less affected in language change and consequently might give an inappropriate picture. We collected data from 140 adult self-reported native speakers of Slovenian (88 female, mean age=37,9, SD=11,4; median age=36) who participated in this experiment voluntarily (indicating online consent), anonymously, and for no material compensation. The participants all spoke the dialects under investigation as indicated in the pre-test demographic questionnaire, respective sample sizes per dialect are shown in the last column of Table 2. All informants had normal or corrected to normal vision and reported no history of neurological disorders. The informants that were not native speakers of selected dialects were excluded from the analysis. We also excluded informants that did not reach a 50% threshold of correctly producing 32 control trials. This led to exclusion of 47 participants. The data from the remaining 93 participants were subjected to analysis.

Table 2: Slovenian dialect with respect to dual preservation.

Dialects	Elicitation SLA	Verb <i>delati</i>	Noun <i>brat</i>	Number of informants
Pohorje mountains (A1)	1955–65	+	+	16
Soča river (A2)	1951	+	+	12
Upper Carniola (A3)	1959	+	+	28
Western (B1)	1954–82	–	+	20
Southern 1 & 2 (B2 & B3)	1957	–	+	17

2.2 Materials

We created 64 colour drawings with a resolution of 300 dpi and size 373x220 pixels (ratio 16:9). Each stimulus stood for a single transitive event involving two characters, the agent and the patient. Three boxes were vertically aligned along

the right margin of the picture; each contained a printed lexeme in standard Slovenian. Since we were not interested in word order, we arranged the lexemes from top to bottom so that they followed the unmarked word order in Slovenian, which is subject-verb-object (SVO): the top and bottom lexemes were nouns (in the nominative singular), and the middle lexeme was a verb (in the infinitive), as shown in Figure 1. Informants were instructed to record a sentence that would best describe the picture using these three lexemes: a verb to name the action, and nouns to name the characters in the event. Informants often replaced the lexemes with dialectal or dialectally pronounced expressions, which suggests that the task actually elicited the dialect rather than a superregional or even the standard language.

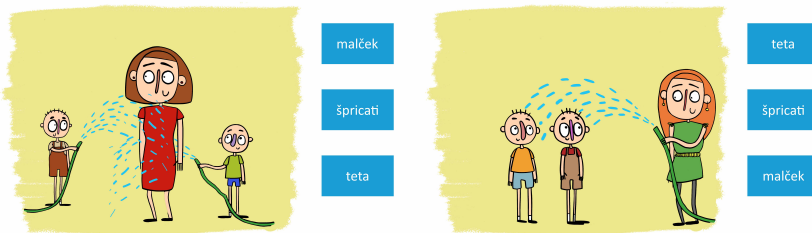


Figure 1: A target set of graphical stimuli.

We prepared 32 target stimuli with 16 different transitive verbs. Each of the 16 verbs from the target set was used twice, always with the same two characters, but with thematic roles reversed (so that each character served once as agent and once as patient). In the boxes, a noun was suggested for each of the characters in the sentence, one was feminine and one masculine. The masculine noun was intended to refer to a character consisting of two entities, i.e., it was intended to elicit a dual form. The feminine noun was intended to refer to a character consisting of either one or three entities, i.e., it was intended to produce a singular or plural form. A single informant had to produce only one sentence out of the set (i.e., half of the target sentences and only one version of each event), for the total of 16 target sentences out of 32. In addition, we prepared 32 control stimuli containing 32 different transitive verbs and 64 different nouns that were counterbalanced for gender ($\frac{1}{2}$ masculine and $\frac{1}{2}$ feminine), number ($\frac{1}{2}$ singular and $\frac{1}{2}$ plural), and sentence function ($\frac{1}{2}$ subject and $\frac{1}{2}$ object). All nouns were repeated exactly once in the control stimuli. Thus, all informants saw 16 target trials and 32 control trials, based on which they recorded 48 SVO-sentences with 48 transitive verbs and 96 gender-balanced nouns.

2.3 Procedure

The experiment was conducted in the online environment Ibex Farm (Drummond 2021), enhanced with the PennController module (Zehr & Schwarz 2018). It generated quasi-random trials for each informant with at least one control trial between the two target trials. The informants conducted the experiment with their own equipment at a location of their choice, but were specifically instructed to do so in a quiet environment and to use their dialect and no formal or standardized language. Prior to the experiment, the informants gave informed consent and completed a brief demographic questionnaire. In the practice session, the informants were instructed to count to five and do a practice item to familiarize themselves with the instructions and stimuli and to learn how to turn the recording on and off. They were then able to play back the recording of the counting and practice item to check the function of the microphone, the volume, and the clarity of their own speech. In the experimental part the informants saw pictures one by one on the computer screen. Each picture appeared on the computer screen at the same time as the keywords. The informants had to put the words together as quickly as possible to form a sentence describing the picture, pronounce the sentence, and record it. The entire experiment lasted between 20 and 25 minutes.

2.4 Transcription, data cleaning and analysis

The recordings were manually transcribed into standard Slovenian, and the grammatical number was coded for each target noun. Prior to statistical analysis we excluded: incomplete recordings (due to premature termination of the recording); incomprehensible recordings or parts of recordings; recordings in which informants used a verb with a non-target sentence structure or kept the verb in the infinitive; recordings in which informants used a noun in a non-target gender, non-target case, non-target thematic role, or non-target sentence function; recordings in which informants retained the target noun in the singular rather than using it in the dual/plural form, consistent with the picture. The remaining 1247 target nouns (i.e., those denoting a character consisting of two entities) were statistically analyzed.

3 Results

3.1 The false plurals on nouns with respect to the dialect

First we checked for number mismatches but did not register a single case in which the subject is dual and the verb is plural – or the subject is plural and the verb is dual. Next, we counted the number of false plurals on nouns in each of the selected dialects to check for consistency between the two dialectal groups (with preserved dual on verbs and with less preserved dual on verbs). The total number of false plurals instead of the expected dual was 152 or 12% of the total data points. Table 3 summarises the results.

Table 3: The frequency of dual and false plural noun forms.

Dialect	more preserved dual on verb						less preserved dual on verb			
	Pohorje (A1)		Soča (A2)		Upper (A3)		Western (B1)		Southern (B2 & B3)	
Dual	216	(95.6%)	90	(75.6%)	375	(94.9%)	220	(83.3%)	194	(79.8%)
Plural	10	(04.4%)	29	(24.4%)	20	(05.1%)	44	(16.7%)	49	(20.2%)

The ratio of false plurals on the nouns in the Western (B1) and Southern (B2 and B3) dialects with less preserved dual on verbs did not differ significantly ($\chi^2(1) = 0.813$, $p = 0.367$). On the other hand, the three dialects with preserved dual on the verb were not homogeneous: the number of false plurals on the nouns in the Soča River dialect were significantly greater than in both the Pohorje ($\chi^2(1) = 28.971$; $p < 0.0001$) and Upper Carniola ($\chi^2(1) = 37.317$; $p < 0.0001$) dialects. The Pohorje and Upper Carniola dialects did not significantly differ in the ratio of false plurals ($\chi^2(1) = 0.026$; $p = 0.87$). These results confirm Jakop's (2008) findings regarding the preservation of noun dual morphology for all dialects except the Soča River dialect. The situation in the Soča River dialect is likely to have changed in 70 years since the data for the Slovenian Linguistic Atlas for the relevant dialects were collected. It should be noted that the Soča River dialect is spoken near the Slovenian-Italian border in a valley that extends into the area where the Western dialect is spoken. We hypothesise that in recent decades it has become easier for speakers of the Soča dialect to commute to neighboring regions and to establish and maintain contacts with speakers of the Italian language and the Western Slovenian dialect, which has led to increased contact and influence of these varieties (with less preserved dual) on their language.

3.2 The effect of the sentential function of a noun

We found a strong correlation between the preservation of verbal dual morphology and the use of dual/plural forms in the target noun with respect to its syntactic position, confirmed statistically by the χ^2 -test and the Cramer coefficient V. For the subject position, Cramer’s V was close to 1, implying a near perfect correlation between noun and verb forms. In contrast, Cramer’s V for the object position was close to 0, indicating an almost complete lack of association between the noun and verb forms. This suggests that the choice of noun form does indeed depend on its syntactic function, as predicted, and that this dependence is strong. The results are summarized in Table 4.

Table 4: The frequency of dual and false plural by sentential function.

Verbal morphology	Elicited noun forms			
	Subject		Object	
	dual	plural	dual	plural
dual preserved	1175	3	1091	87
dual less preserved	7	62	66	3
$\chi^2(1)$ -test	1041.1		0.50	
p-value	<0.0001		0.47	
Cramer’s V	0.9216		0.0268	

To better understand the difference in sentential function, we evaluated this correlation separately for dialects with the preserved dual on the verb and dialects with the less preserved dual on the verb, as shown in Table 5.⁵ In the dialect with the preserved verbal dual morphology, the difference between the number of plural forms instead of dual forms in the subject and in the object is statistically significant, while in the dialect with less preserved verbal dual morphology, the difference between the number of plural forms instead of dual forms in the subject and in the object is not statistically significant. As for the subject position, the difference between the number of plural forms instead of dual forms in the dialects with preserved dual and less preserved dual is statistically significant, while for the object position, the difference between the number of plural forms instead of dual forms in the dialects with preserved dual and less preserved dual is not statistically significant.

⁵The first number in the cell shows the number of occurrences of the false plural compared to the total number of target nouns (second number). The third number is the percentage.

Table 5: Analyzed target nouns by sentence function and dialect group.

			noun			
			subject	object		
verb	+dual	p<0.0001	15/353 (04%)	45/386 (12%)	p>0.05	
	-dual		50/260 (19%)	42/248 (17%)		
			p<0.025			
			p>0.05			

Finally, we modelled the results. We added the target response (the dual of a noun referring to two entities) to the model as a reference value and used it to estimate the probability of the non-target response (the plural of a noun referring to two entities). Because the outcome was a categorical variable (either dual [1] or plural [0]), we used a mixed-effects logistic linear model (Jaeger 2008, Winter 2020) for binary outcomes via the *glmer* function in the *lme4* package in version 4.0.2 of the open-source R computing environment (R Core Team 2020) to test for dependence on a linear combination of independent predictor variables, while accounting for possible random noise. The independent variables in our model were syntactic position (noun in subject or object position) and dialect type (dialects with preserved dual on verb or dialects with fewer preserved dual on verb). Since we could not rule out the possibility that the informants' responses to the various conditions depended on the dialect they spoke, we also included in the model a test for the interaction between syntactic position and dialect type. The quality of the statistical model (i.e., the degree of agreement between the measured values and the values expected within the model) was validated using the Akaike information criterion: Compared with alternative models containing the same independent and dependent variables, the model with the selected function fitted our measured values best. Confidence intervals (CI) and p-values were calculated using the Wald test. The model results are in Table 6.

The statistically significant odds ratio (OR) for a non-target response in the subject position in the noun condition for dialect group with preserved dual of the verb is 0.01, which corresponds to a probability of 1%. The odds ratio for a non-target response in the subject position in the noun condition compared to the object position for both dialect groups combined is 6.31, which was found to be not statistically significant. At the same time, the effect of dialect group is statistically significant, as the odds ratio for a non-target response for dialects with a preserved dual verbal morphology (in both syntactic positions) is 0.11. This means that a non-target response is less likely for dialects with a preserved dual

Table 6: Statistical model.

Factors	OR	CI	p
(intercept)	0.01	0.00–00.07	<0.001
Function [object]	6.31	0.69–57.33	0.102
Group [+dual on verb]	0.11	0.01–00.99	0.049
Function [object] * Group [+dual on verb]	10.11	1.33–76.60	0.025
Number of informants		93	
Number of stimuli		32	
Data points (target nouns)		1247	

verbal morphology than for dialects with a less preserved dual verbal morphology (10% vs. 90% odds ratio for a non-target response). The interaction of dialect group and syntactic position is also statistically significant: a likelihood ratio of 10.11 shows that non-target responses are statistically significantly more likely to occur for object position compared to subject position in dialects with preserved verb dual morphology, but not in dialects with less preserved verb dual morphology. This difference is also evident in Figure 2.

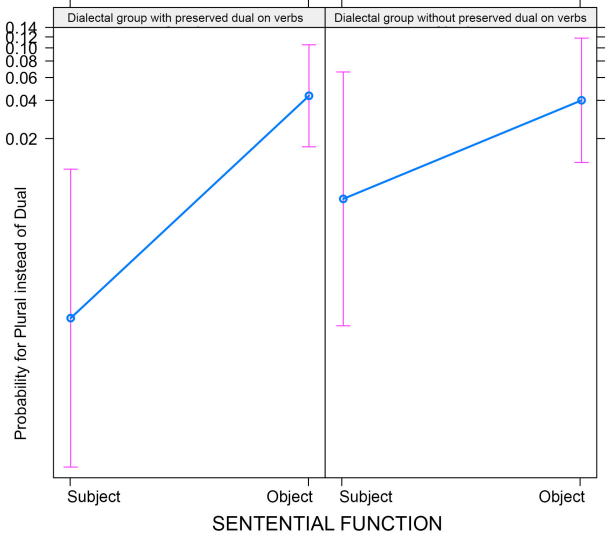


Figure 2: Comparison of the probability of false plural when syntactic position is crossed with dialect type.

To further investigate this interaction, we performed pairwise comparisons with the *emmeans* package in R, taking into account Tukey's adjustments for multiple comparisons. These comparisons reconfirmed that in dialects with preserved dual morphology of the verb, the probability of a target response is statistically significantly higher for a noun in subject position than for a noun in object position ($\beta = -4.15$, $SE = 1.37$, $z = -3.03$, $p = 0.01$). In other words, informants were more likely to use a dual form for the target noun in subject position than for the target noun in object position. This was not the case in dialects with a less preserved dual morphology of the verb, where the likelihood of a dual form did not differ significantly with respect to the syntactic position of the noun ($\beta = 1.84$, $SE = 1.12$, $z = 1.63$, $p = 0.36$). Thus, informants in this dialect group were equally likely to substitute the dual for plural in any syntactic position.

4 Discussion

In this sentence production experiment the realization of the dual in a noun was related to the sentential function of this noun. In dialects with preserved dual verb forms, the use of a noun in the dual to refer to two entities depends on its syntactic position (in subject position it is higher than in object position). This is not the case in dialects with less preserved dual verb forms. On this basis, we can both predict and explain why pluralization of the dual in Slovenian dialects occurs first in contexts where the noun is not morphologically marked for the dual (in the case of coordinated noun phrases or silent personal pronouns).

It should be noted that the pattern of subject-verb agreement is extremely strong in our sample, as we did not register a single case in which the subject is dual and the verb is plural – or the subject is plural and the verb is dual. Thus, when the dual is dropped in favor of the plural, it is dropped from both the subject and the verb. Second, we noted that in the dialects that, according to Jakop (2008) and based on the Slovenian Language Atlas, have lost the dual morphology in the verb, the dual is still used in the verbs, although less than in other dialects.

Next, considering that the dual forms of nominative and accusative masculine are homophonous, one might expect a “transitive” supporting effect of subject-verb agreement, in the sense that greater use of the dual in the nominative (due to agreement with the verb) would lead to greater use of the dual in the accusative (which is identical in form to the nominative). Our results show no such effect and suggest that the masculine nominative and accusative forms in the dual, although identical, are treated morphologically independently.

Finally, the results are interesting from a theoretical point of view. How can the diachronic changes (i.e., loss of the dual) in certain Slovenian dialects be explained in such a framework, especially when at the same time mismatch or split agreement (where the verb and the subject carry different features) are not supported? We tentatively hypothesise that homophony is an intermediate step in the gradual loss of the dual. First, the plural form is reinterpreted to refer not only to “more than two entities” but also to “more than one entity”.⁶ At this stage, dedicated dual forms may coexist with neutralized plural forms until the neutralized forms predominate and, as a result, the paradigm becomes partially homophonous until, finally, the dual forms are completely abandoned in favor of the plural forms. Under the assumption that diachronic change follows a principle of Economy (e.g. Martinet 1955), the “more than one entity” interpretation of the plural wins over the “more than two entities” interpretation since the former refers to a superset of possible referents in comparison to the latter and thus has a wider descriptive potential in comparison to the latter. According to this hypothesis, either the noun or the verb may be neutralized first, but the not-yet-neutralized form would support the dedicated interpretation of the homophonous form, even if the forms are no longer distinct.

5 Conclusion

Using a picture description task based on three given key words, we tested if the actual use of the dual form of a noun is related to the preservation of dual forms in Slovenian dialects and the sentential function of that noun. There was no split-matching (i.e., plural subject and dual verb or vice versa) in the sample studied. The results showed that speakers use the dual form of the noun in subject position more often than in object position. In the dialects with better preserved dual forms of the verb, the number of non-target plural nouns in subject position was lower than in object position because this dialect group preserved the dual forms of the verb and verb agreement seems to play a supporting role in preserving the dual morphology of the subject. In the dialect group where the dual forms of the verb were less preserved, the number of dual forms of the noun did not depend on its syntactic position because the dual morphology of the verb was lost.

⁶This step can conceivably be couched within the pragmatic theory of plural (cf. Sauerland 2008, Spector 2007), according to which the pragmatically enriched interpretation of plural results from enriching its lexical meaning associated with one or many objects with relevant pragmatic inferences. In languages featuring singular and plural number marking, this is an anti-singularity inference. In languages featuring singular, dual, and plural number marking, both an anti-singularity and anti-duality inferences strengthen the meaning of the plural.

Acknowledgments

The research was financially supported by the ARIS research project “Linguistic transfer in the pragmatic domain: Slovenian speakers in a multilingual environment” no. J6-2580 (PI: Penka Stateva).

Abbreviations

SG	singular	NOM	nominative
DU	dual	ACC	accusative
PL	plural	SVO	subject-verb-object
M	masculine	SLA	Slovenian Linguistic Atlas
F	feminine	CI	confidence interval
3	third person	OR	odds ratio

References

- Benedik, Francka. 1999. *Vodnik po zbirki narečnega gradiva za slovenski lingvistični atlas (SLA)*. Ljubljana: Založba ZRC.
- Berdicevskis, Aleksandrs & Arturs Semenuks. 2022. Imperfect language learning reduces morphological overspecification: Experimental evidence. *PLoS ONE* 17(1). e0262876. DOI: 10.1371/journal.pone.0262876.
- Corbett, Greville G. 2000. *Number*. Cambridge: Cambridge University Press.
- Derganc, Aleksandra. 2006. Nekatere značilnosti dvojine v slovenščini. *Slavistična revija* 54(1). 415–434.
- Drummond, Alex. 2021. *Ibex farm*. <https://spellout.net/ibexfarm>.
- Greenberg, Joseph H. 1966. *Language universals*. Mouton The Hague.
- Ivanov, V. 1983. *Istoričeskaja grammatika russkogo jazyka*. Moscow: Prosveščenie.
- Jaeger, T. Florian. 2008. Categorical data analysis: Away from ANOVAs (transformation or not) and towards Logit Mixed Models. *J Mem Lang*. 59(4). 434–446. DOI: 10.1016/j.jml.2007.11.007.
- Jakop, Tjaša. 2008. *Dvojina v slovenskih narečjih*. Ljubljana: Založba ZRC.
- Kenda-Jež, Karmen. 2002. Model idealnega govorca v slovenskih dialektoloških raziskavah. In Marko Jesenšek, Bernard Rajh & Zinka Zorko (eds.), *Med dialektologijo in zgodovino slovenskega jezika*, 150–165. Maribor: Slavistično društvo. <https://press.um.si/index.php/ump/catalog/book/41>.
- Martinet, André. 1955. *Économie des changements phonétiques: Traité de phonologie diachronique*. Bern: A. Francke.

- Nevins, Andrew. 2011. Marked targets versus marked triggers and impoverishment of the dual. *Linguistic Inquiry* 42(3). 413–444. DOI: 10.1162/LING_a_00052.
- Orel, Irena. 2019. Ženske dvojske glagolske oblike v starejšem slovenskem knjižnem jeziku. *Slavistična Revija* 67(2). 273–280. <https://srl.si/ojs/srl/article/view/2019-2-1-15>.
- Polinsky, Maria. 2018. *Heritage languages and their speakers*. Cambridge: University Press.
- R Core Team. 2020. R: A language and environment for statistical computing. Vienna: R Foundation for Statistical Computing. <https://www.R-project.org/>.
- Ravid, David & Lubna Hayek. 2003. Learning about different ways of expressing number in the development of Palestinian Arabic. *First Language* 23(1). 41–63. DOI: 10.1177/0142723703023001003.
- Sauerland, Uli. 2008. Implicated presuppositions. In Anita Steube (ed.), *Sentence and context: Language, context, and cognition*, 581–600. Berlin: Mouton de Gruyter. DOI: 10.1515/9783110209303.4.581.
- Slobodchikoff, Tatyana G. 2019. *The evolution of the Slavic dual: A biolinguistic perspective*. Lanham: Rowman & Littlefield.
- Spector, Benjamin. 2007. Aspects of the pragmatics of plural morphology: On higher-order implicatures. In Uli Sauerland & Penka Stateva (eds.), *Presupposition and implicature in compositional semantics*, 243–281. London: Palgrave Macmillan. DOI: 10.1057/9780230210752_9.
- Stepanov, Arthur & Penka Stateva. 2018. Countability, agreement and the loss of the dual in Russian. *Journal of Linguistics* 54(4). 779–821. DOI: 10.1017/S0022226718000130.
- Tesnière, Lucien. 1925. *Les formes du duel en slovène*. Paris: Ancienne Honoré.
- Toporišič, Jože. 2000. *Slovenska slovnica*. Maribor: Obzorja.
- Winter, Bodo. 2020. *Statistics for linguists: An introduction using R*. New York: Routledge.
- Zehr, Jérémy & Florian Schwarz. 2018. *PennController for internet based experiments (IBEX)*. <https://www.pcibex.net/>.