

# Interlanguage correspondences and their manifestation in receptive bilinguals



Adam Kříž (Charles University, Prague) –  
Jan Chromý (Charles University, Prague)

## ABSTRACT

The current paper presents the results of a qualitative analysis of speech errors made during the L2 picture-naming task in the context of two mutually intelligible languages. The task was performed by native speakers of Slovak and its target language was Czech. The errors demonstrate that production is partly based on the so-called interlanguage analogies and interiorized rules derived from them. Such issues have been rather overlooked in the previous literature. Our qualitative findings offer a viable starting point for formulating novel hypotheses for quantitative studies on lexical access in bilinguals. The evidence discussed here shows that examining less commonly studied languages in their specific contexts can substantially enrich the research that has traditionally focused on better studied languages, such as English, French, German or Dutch.

## KEYWORDS

bilingualism, Czech, picture-naming task, production errors, Slovak

## DOI

<https://doi.org/10.14712/18059635.2022.1.6>

## 1 INTRODUCTION

One of the central questions in the (psycholinguistic) research on bilingualism is to what extent both are languages activated during language comprehension and production. The existing scientific evidence points clearly to the conclusion that the cross-language activation in a bilingual is always present, even when processing items from just one of the languages available to the bilingual individual (e.g., de Groot 2011; Broersma, Carter and Acheson 2016; Frances, Navarra-Barindelli and Martin 2021; Hoshino, Beatty-Martínez, Navarro-Torres and Kroll 2021).

However, it is still debatable how items from two different languages really interact with one another and at what level of language processing the interaction takes place. The current paper<sup>1</sup> deals with several important theoretical accounts of bilingual language processing, with a specific focus on speech production and cognates. The research tradition in this field has largely focused on examining generally well-studied languages, such as English, French, Spanish or Dutch. In contrast, the current paper presents the results of a qualitative analysis of picture-naming data

---

1 The work was supported by the European Regional Development Fund project “Creativity and Adaptability as Conditions of the Success of Europe in an Interrelated World” (reg. no.: CZ.02.1.01/0.0/0.0/16\_019/0000734).



from an empirical study conducted on rather understudied languages, namely Czech and Slovak. Specifically, we examine the types of errors made by native speakers of Slovak in naming pictures in Czech. The investigation of errors in a picture-naming task in two closely related, mutually intelligible languages represents a unique research design which can consequently provide important findings about bilingual lexical access in general.

## 2 COGNATES AND BILINGUAL LANGUAGE PROCESSING

Generally, the evidence for cross-language activation comes mainly from studies manipulating lexical stimuli with formal similarities across the languages under investigation. Among these, the so-called cognates occupy a prominent position.

Cognates are defined as words which share meaning and form across the languages. The formal overlap can be total (identical cognates, e.g., *finger* in English, *Finger* in German) or just partial to various degrees (e.g., *garden* in English, *Garten* in German). Cognates fall into what Woolard (1999) terms bivalent elements. As she shows, mainly in the Castilian-Catalan setting, such elements can be deployed socially meaningfully and strategically by members of the respective speech communities. At the same time, it can be argued that they serve as tools for preventing a misunderstanding in bilingual communication (see Nábělková and Sloboda 2008). Thus, as such, they play a significant role in language adaptation and accommodation. The reason to believe that they can be utilized strategically and intentionally stems also from findings that they are learned more easily than non-cognate forms (e.g., Comesaña, Soares and Lima 2010) and are processed/produced more quickly and accurately than non-cognates (e.g., Lemhöfer, Huestegge and Mulder 2018).

These findings are known as the cognate facilitation effect. The cognate facilitation effect in the simple picture-naming task can be explained by the double activation boost of the phonological form of the cognate. The boost comes from two languages, instead of only one in the case of non-cognate (e.g., Kroll, Bobb, Misra and Guo 2008: 421). Costa, Santesteban and Caño (2005: 99–101) postulate that lexical nodes of co-activated units trigger the activation of corresponding phonological units. Because cognates share (more or less) the same form, the phonological units receive a relatively high activation. Furthermore, according to the authors, the phonological segments send the activation back to the formerly activated lexical units and they, in turn, send it to the formal segments. Several iterations are assumed until the target representation is selected. Apparently, the presented view is based on the idea of interactive processing, involving feedback between the respective levels.

This view accounts for the difference in the effect magnitude between the dominant and non-dominant languages. The activation level of the lexical nodes belonging to the dominant language is higher, hence they can effectively boost the activation of the segments of a non-dominant language target. In contrast, when producing words in the dominant language, their activation is already high enough, so any additional boost is not necessary or effective (although see Klaus, Lemhöfer and Schriefers 2018,

who demonstrate using a picture-word interference task with L1 Dutch and L2 English speakers that even L2 properties affect L1 production).

The proposed existence of the feedback between lexical and phonological nodes, as incorporated in Costa, Santesteban and Caño (2005), can also mean competition at the lemma/lexical level. Broersma, Carter and Acheson (2016) argue that this competition can lead to the cognate inhibition effect. At the same time, they add that the inhibition can be outweighed by the benefits at the phonological level. Additionally, Li and Gollan (2018: 926) consider the higher cross-language competition in cognates as a precondition for higher intrusion and self-correction rates in cognates, as opposed to non-cognates. The authors (*ibid.*) frame this consideration within the conflict-monitoring hypothesis concerning the self-monitoring of speech (Nozari, Dell and Schwartz 2011).

The above accounts of cognate effects suppose that the effect is the result of the general features of the processing system. However, there are also different approaches based on the postulation that cognates are represented differently from non-cognates. Costa, Santesteban and Caño (2005: 97–98) mention that the effect can stem either from larger morphological overlap (i.e. shared morphological representations) or from larger conceptual overlap (i.e. shared conceptual representations). Nevertheless, the authors reject both approaches. First, they argue it is unclear how the final correct pronunciation of the cognates would be executed if the stem were shared across the languages. According to Costa, Santesteban and Caño (2005), the assumption that there are rules for form transformation between languages seems to be rather unconvincing. They claim that there is no obvious systematicity in the creation of cognates from the same stem across languages. Second, the same authors claim it is unclear to them how the formal similarities can impact semantic dimensions. However, De Groot (2011: 247) offers a possible explanation. She references studies on L2 vocabulary acquisition and argues that during this process, L2 learners can perceive cognate similarity as a cue, saying that the meaning of the form(s) is more similar than it is the case in other words. As a result, the different meaning representations in cognates can arise.

One link between the processing and representational aspects of bilingual production can be found in the concepts of language cue and language tag. The WEAVER++ model (Roelofs 2018; Dell and Cholin 2012), which is a specific modification of Levelt's model (Levelt 1989), suggests that the intention to speak a particular language is encoded in the conceptualizer. The intention is materialized in the form of a language cue as part of the preverbal message. In addition, the model holds that each lemma encompasses the information on language membership in the form of a language tag. The match between the language cue and the language tag leads to a higher activation of the target and consequently to its production (see de Groot 2011: 228–229). De Groot (*ibid.*: 229) adds that concepts that are similar to the language cue and tag are broadly accepted in the models of bilingual speech production, indicating that they help bilingual individuals to identify the intended language during language processing.

The task of language cue/tag can be interpreted differently depending on which theoretical stance is taken. Inherent in most models is the assumption that the activation going from the conceptual level to the lexical level is generalizable across



languages (Klaus, Lemhöfer and Schriefers 2018: 867). The real inconsistencies and divisions can be found in the discussions about the language-specificity of the lexical selection (*ibid.*: 867).

Broersma, Carter and Acheson (2016) survey the views that treat the competition for selection as an optional mechanism or one that emerges mainly in less proficient bilinguals (e.g., Costa and Santesteban 2004). This reflects the claims made by Kroll, Bobb, Misra and Guo (2008: 417) that “the intention to speak one language alone may be influenced by the relative dominance of the two languages for the bilingual, by the context in which spoken production occurs, and by features of the two languages themselves”. This topic was discussed also by de Groot (2011: 229), among others, who elaborates on the theoretical approaches that focus on tackling the communicative situations, whereby language mixing does not pose any serious issue for a speaker (*ibid.*). She recalls De Bot’s approach (De Bot 1992), which assumes the simultaneous generation of two speech plans, which allows rapid language switches. Additionally, de Groot cites De Bot and Schreuder’s approach (De Bot and Schreuder 1993) according to which the language cue is weighted, depending on the consequences of language mixing in a particular situation. It holds that in contexts where language mixing is not uncommon and communicatively disruptive, the weight assigned to the language cue can be low, which in turn allows the production of a mixed output.

### 3 PRESENT STUDY: THE CZECH-SLOVAK CONTEXT

The psycholinguistic empirical research addressing the processing and production issues in bilinguals covers studies from various settings and contexts. A crucial difference lies in the targeted population. The focus has been on both monolingual and bi-/multilingual populations (Antón-Méndez and Gollan 2010), on proficient and non-proficient bilinguals (Allen and Conklin 2013), on speakers acquiring their second language in the natural environment and on speakers learning L2 in classroom settings (Poarch and van Hell 2012), on clinical and non-clinical populations (Grasso, Peña, Bedore, Hixon and Griffin 2018), on speakers using two related languages (Dutch-English; Muylle, Van Assche and Hartsuiker 2022), two relatively distant languages (Spanish-English; Hoshino and Kroll 2008) or totally different languages with different orthographic systems (Japanese-English; *ibid.*). However, there are still language constellations which are understudied and which have a potential to contribute to the general debates about the essence of bilingual production and processing. One such constellation is the Czech-Slovak language context.

Both languages belonging to the West-Slavic language family are genetically and socially closely related, which underlies their mutual intelligibility. Structurally, Czech and Slovak can be perceived as a dialect continuum (Dickins 2009). As such, shared cognates (be it identical or non-identical ones) dominate over non-cognates in both lexicons. This can be seen as one peculiarity characterizing the languages in question. Word forms that share meanings but look and sound completely different are salient exceptions and often become subjects of popular comments or jokes. Re-

garding formally shared items, Nábělková and Sloboda (2008) distinguish between bivalent elements (see Woolard 1999 above) and parallel elements. The former correspond to identical cognates, the latter to non-identical cognates. A considerable number of parallel elements contain systematic interlanguage correspondences. Here, Sloboda's (2004) term of interlanguage analogies is of relevance to understanding how non-identical cognates function. Sloboda (ibid.) claims that Czech and Slovak speakers develop interlanguage analogies to identify the units of a second language. Analogies can cover more language levels and are derived from monitoring regularities in the differences between languages, e.g. the awareness of the existence of non-identical cognate pairs, such as *smích* (Czech) — *smiech* (Slovak) ‚laugh‘, *líh* — *lieh* ‚alcohol‘, *mícha* — *miecha* ‚spinal cord‘, can lead to forming an analogy *í* — *ie*, i.e. an interiorized rule, which says that *í* (in certain positions and certain words) corresponds to *ie* in the second language.

The Czech-Slovak sociolinguistic context resembles the situation in Scandinavian languages. Thus certain terms describing its core features are adopted from or, conversely, can be applied to Swedish-Danish-Norwegian sociolinguistic situations. One such term is semi-communication (Haugen 1966). The term was introduced specifically for inter-Scandinavian communication and stresses the fact that divergence between languages can cause communication noise and misunderstandings. Over the course of time, its popularity among researchers has decreased as it is perceived to be misleading, and alternative terms are used instead (Braunmüller 2002). Receptive bilingualism can be taken as an example (e.g., Zeevaert and ten Thije 2007). The term refers simply to communication situations in which a speaker sticks to his/her own native language, while being able to understand the second language used by the interlocutor.

Despite the mutual comprehensibility, the switching to the second language or attempts to acquire it can be detected, especially in the context of living/being in the second language environment. As far as Slovaks living in the Czech Republic are concerned, the accommodation to Czech is supported by the publicly discussed problems of Czechs — and namely young Czechs — to understand the Slovak language properly (Sloboda 2004). The accommodation to Czech is also driven by tendencies to adjust behaviors to the majority norms (ibid.) or is based on Slovak speakers perceiving the vitality of the Slovak language as low (Sloboda 2005). It is worth noting that the decision to acquire Czech is usually not accompanied by attempts to do so officially, in classes or by using grammar books. The second language acquisition in this context takes place naturally by observing patterns that occur in everyday speech. The data suggest that the Slovaks' accommodation to Czech concerns mainly non-cognates (Musilová 2006).

## 4 PRESENT STUDY: METHODOLOGY

### BACKGROUND AND GOALS

We conducted an empirical study on two subpopulations of native speakers of Slovak using one of the most widely used paradigms in speech production research, namely





L2 picture-naming task. Two subpopulations were tested — Slovaks living in Slovakia and Slovaks living in Prague on a long-term basis. Thus we tested a) receptive bilinguals in their L1 environment and b) receptive bilinguals immersed fully in the L2 environment.

The research that aims at speech production (in bilinguals) using behavioral methods is concerned mainly with the analysis of reaction times in the tasks employed. It is not to say that the research on bilingualism is totally oblivious to the (qualitative) analysis of errors, but those are mostly tracked in studies with clinical populations. For example, Gollan, Stasenko, Li and Salmon (2017) focus on reading errors in Spanish-English bilinguals with Alzheimer's disease, and McMillen, Griffin, Peña, Bedore and Oppenheim (2020) present work on error patterns in a bloc cyclic naming task in Spanish-English bilingual children with and without developmental language disorder.

Existing research distinguishes errors like intrusions, semantic errors, self-corrections, no responses, cut-offs or phonological errors including alliterations or non-words (McMillen, Griffin, Peña, Bedore and Oppenheim 2020). Poulisse (2000) classifies several types of speech errors (slips of the tongue) and discusses their origin. In her paper on Dutch-English bilinguals and their speech, she puts a special emphasis on the so-called L1/L2 blends consisting of features from both languages. See, for instance, *he cwame* as a mixture of the Dutch *kwam* and the English *came*. She interprets such blends as the result of simultaneous activation of both languages. Any comments on, or definition of, systematic errors mirroring interlanguage correspondences are absent in her list.

Our main goal is to show that in the receptive bilingual situation with two genetically closely related languages the L2 word production can lead to specific incomplete, non-target units. These units can be described as L1/L2 blends containing features of both languages in question. We believe they are based on what Sloboda (2004) calls interlanguage analogies. The occurrence of such forms can stem from blurred/absent representations of the L2 word form itself. The blurred/absent representations are then the result of the speakers' lower need to fully, intentionally acquire L2 material and consequently of their relatively lower L2 proficiency.

## PARTICIPANTS

The group of Slovak speakers living in Slovakia comprised 30 participants (16 women and 14 men). Their mean age was 20.8 years (range = 18–24, SD = 1.5). The group of Slovak speakers living in the Czech Republic comprised 46 participants (31 women, 15 men) with a mean age of 33 years (range = 20–60, SD = 9.3), the mean age of migration to the Czech Republic was 22.7 years (range = 17–47, SD = 6.5) and the mean length of residence in the Czech Republic was 9.9 years (range = 20 months — 37 years, SD = 7.2).

Several self-assessment scales were used to measure the participants' proficiency in Czech or Slovak and their exposure to these languages. Their responses are summarized in Table 1; the numbers indicate how many respondents picked the corresponding option on the scale.



Participant group	Variable and points on the scale					
	Czech proficiency					
	very bad	bad	average	good	very good	native-like
Slovaks in Slovakia			12	14	4	
Slovaks in Czechia			2	16	25	3
	Czech spoken production					
	almost never		rarely	sometimes	often	very often
Slovaks in Slovakia	16		11	3		
Slovaks in Czechia	5		6	13	11	11
	Exposure to Czech audio-visual content					
	almost never		rarely	sometimes	often	very often
Slovaks in Slovakia	1		3	4	11	11
Slovaks in Czechia			5	11	15	15
	Slovak spoken production					
	almost never		rarely	sometimes	often	very often
Slovaks in Czechia	1		1	5	10	29
	Exposure to Slovak audio-visual content					
	almost never		rarely	sometimes	often	very often
Slovaks in Czechia	9		14	13	7	3

**TABLE 1.** Self-assessment scales

The table shows that the proficiency and production of Czech is at a higher level in the Slovaks in Czechia than in the Slovaks in Slovakia, while the passive encounter does not differ much.

No participant reported any clinical speech or cognitive deficit or vision problems. Just two participants (both from the group of Slovaks living in Slovakia) stated that their native language, along with Slovak, is Hungarian. The remaining participants all identified themselves as monolingual Slovak speakers (note that in the Czech-Slovak sociolinguistic context, it is very rare that Czech/Slovak speakers consider themselves bilingual, even though they understand the second respective language, see Nábělková 2008). Twenty-five Slovaks in Czechia assessed their current Slovak competence as worse than before moving to the country. The other participants in the group see no change.

## STIMULI

For the stimuli selection, the BOSS database was used (Brodeur, Dionne-Dostie, Montreuil and Lepage 2010; Brodeur, Guérard and Bouras 2014). The database contains color photographs of various objects on a white background. 156 pictures were selected, covering several semantic categories (vegetables and fruits, food, animals, tools, buildings, products, body parts, and furniture) and names classified as identi-



cal cognates ( $n = 62$ ), non-identical cognates ( $n = 58$ ) and completely different words ( $n = 36$ ). Non-identical cognates included words differing just by one phonological unit (*kohout* /Czech/ — *kohút* /Slovak/ ‚cockerel‘) or more elements (*loket* /Czech/ — *lakeť* /Slovak/ ‚elbow‘). Just 136 items were the subject of analysis as some of the items served as fillers.

## PROCEDURE

The participants took part in individual experimental sessions. Written instructions were presented in the participants' native language for the Slovaks living in Slovakia, whilst the Slovaks living in the Czech Republic could choose between Slovak or Czech instructions. When explaining the instructions and communicating with the participants, the experimenter used the Czech language. The participants gave their informed consent and got a financial reward for their participation.

The task was designed using the DMDX software package (Forster and Forster 2003) and was run on a Dell Inspiron 15 5559 laptop, with a monitor size of 15,6" and resolution of 1920x1080. Each individual trial consisted of a presentation of a fixation cross (+) for 350ms, which was then replaced by a photograph. The photograph remained on the screen for the maximum of 500ms. The trials were presented in a random order for each participant. The participants wore headphones with a microphone and were asked to name the object on the screen as quickly as possible. They were told to produce a Czech one-word name. The recordings were coded and analyzed using the software package CheckVocal (Protopapas 2007).

After the task, each participant completed a short interview in which they were asked questions about their own task performance.

## 5 PRESENT STUDY: RESULTS AND DISCUSSION

The participants in both groups produced various types of errors ranging from non-responses to total intrusions (e.g., saying a Slovak word, non-cognate or non-identical cognate, instead of a Czech form). Our focus was on a specific speech error type. We assumed that the close relatedness of the languages can lead to systematic L1/L2 blends reflecting frequent interlanguage correspondences/rules/analogies (we use all three terms interchangeably). Crucially, we detected such blends in both participant groups.

In what follows, we provide examples and a short classification of the blends. As our intention is merely to present evidence for this error type, we do not make any proper distinction between the participant groups when summarizing the data. Only at some places we mention whether a given error comes from a Slovak living in Slovakia or in Czechia. It is done just to underscore the fact that both groups exhibit similar behaviors.

Subsequently, we turn our attention also to intrusions as we think they are indicative of certain aspects of language processing which hold also for L1/L2 blends. Finally, we try to identify primary mechanisms underlying the speech error types under analysis. Generally, we claim that incorrect speech outputs stem either from the specifics of L2 word representation or from issues in the speech production (or both).



## L1/L2 BLENDS WITH INTERLANGUAGE CORRESPONDENCES

This type of errors (blends) supports the notion of independent representations of regular interlanguage analogies. To give a general example of an interlanguage analogy, Slovak ending *-ok* corresponds in many Czech equivalents to *-ek*, for instance, *lepok* (Slovak) — *lepek* (Czech) ,gluten', *korok* (Slovak) — *korek* (Czech) ,cork'. Both languages contain also equivalents consisting of identical sounds (in case of *-ok* — *-ek* it is the identical cognates *potok* ,brook' or *oblek* ,suit', for example). However, we propose that the number of pairs containing the *-ok* — *-ek* distinction and the frequency of their use make the distinction salient enough to be a basis for the creation of an interlanguage rule/analogy, as is suggested by Sloboda (2004). Such analogies are then represented as special psycholinguistic units in the speakers' minds and can be a part of language processing/production (similarly to the representations of phonemes, affixes etc.). Depending on several factors (general language activation, language proficiency, word activation, speech settings, etc.), a speaker can get confused and apply an analogy inappropriately (e.g. in a word which is identical across the languages).

In the picture-naming data from our experiment, we identified several types of non-existing words which differ from the targets only by a wrongly applied analogy. The inappropriate application of an analogy included the insertion of a Czech version of an analogy: 1. into an identical cognate, 2. into a non-identical cognate, 3. into a Slovak non-cognate, 4. into a Czech non-cognate, further 5. placing a Slovak version of an analogy into a Czech non-cognate and 6. combining Czech and Slovak versions of analogies in non-identical cognates.

Regarding type 1 (Czech version of an analogy in an identical cognate), we observed a participant (a Slovak in Czechia) producing the non-word *žralok* which was immediately repaired as *žralok* (see the analogy *-ok* — *-ek* discussed above; the word *žralok* ,shark' is an identical cognate). A similar outcome occurred also in a participant (a Slovak in Czechia) naming the picture of an elephant *slůn*, even though both languages have the same name for the animal, *slon*. The non-word was immediately changed to the correct word by the subject. This time, the initial mismatch was apparently derived from the observation of the *ó/o/ô* (Slovak) — *ů* (Czech) correspondence, e.g., *trón* (Slovak) — *trůn* (Czech) ,throne', *sol'* (Slovak) — *sůl* (Czech) ,salt', *stól* (Slovak) — *stůl* (Czech) ,table'. Other examples include forms like *tulen* for seal (produced by three Slovaks in Slovakia and six Slovaks in the Czech Republic, the correct word is *tuleň*, based on the analogy *-eň* — *-en*, e.g. Slovak word: *jeleň*, Czech word: *jelen* ,deer'); or *soud* (produced by two Slovaks in Slovakia, once corrected to the correct form *sud* ,barrel', and produced by two Slovaks in the Czech Republic, based on the analogy *u/ů* — *ou*, e.g. Slovak word: *kút*, Czech word: *kout* ,corner').

Type 2 (Czech version of an analogy in a non-identical cognate) covers outputs like *studně* (produced by two Slovaks in Slovakia and four Slovaks in Czechia, the analogy is *a* — *e* at the end of words following certain consonants, cf. Slovak word *studňa*, Czech word *studna* ,water well') or *kuš* (produced by ten Slovaks in Slovakia, motivated by the analogy *ô* — *ů*, cf. Slovak word *kôš*, Czech word *koš* ,bin').

Examples of type 3 (Czech version of an analogy in a Slovak non-cognate) include *dumbír* (produced by a Slovak in Slovakia; analogy *ie* — *í*, Slovak word *dumbier*, Czech word *zázvor* ,ginger'), *lívik* (produced by a Slovak in Slovakia, then corrected to *lievik*;





produced also by a Slovak in the Czech Republic, analogy *ie* — *í*, Slovak word *lievik*, Czech word *trychtýř* ‚funnel‘), *ceruze* (produced by a Slovak in Slovakia and a Slovak in the Czech Republic; analogy *a* — *e* at the end of words following certain consonants, Slovak word *ceruza*, Czech word *tužka* ‚pencil‘) or *poupava* (produced by three Slovaks in Slovakia and by four Slovaks in Czechia, analogy *ú* — *ou*, Slovak word *púpava*, Czech word *pampeliška* ‚dandelion‘).

Type 4 (Czech version of an analogy in a Czech non-cognate) is represented by non-words like *berouš[ka]* (produced by a Slovak in Czechia, then corrected to the intended Czech form *beruška* ‚ladybug‘, analogy *u* — *ou*, Slovak word *lienka*) and *měruňka* (produced by a Slovak in Slovakia, analogy *me* — *mě*, Slovak word *marhuľa*, Czech word *meruňka* ‚apricot‘).

Type 5 (Slovak version of an analogy in a Czech non-cognate) involves instances such as *brusla* (produced by two Slovaks in Slovakia, analogy *a* — *e* at the end of words following certain consonants, Slovak word *korčuľa*, Czech word *brusle* ‚skate‘) or *židla* (produced by two Slovaks in Slovakia, analogy *a* — *e* at the end of words following certain consonants, Slovak word *stolička*, Czech word *židle* ‚chair‘).

Type 6 (combination of Czech and Slovak versions of analogies in non-identical cognates) is very heterogenous. We can name two examples: *goule* (in total, including corrections, the form *goule* was produced by three Slovaks in Slovakia and by two Slovaks in Czechia) and *netopýř* (produced by five Slovaks in Slovakia and by seven Slovaks in Czechia). The participants produced the form *goule* when naming the picture whose correct (Czech) target is *koule* ‚ball‘. The Slovak equivalent sounds *guľa*. The cognate involves three interlanguage analogies: *u* — *ou* in the middle, *a* — *e* at the end and *g* — *k* at the beginning. Thus, the latter analogy was produced in its Slovak version, the other two in the Czech one. One Slovak in the Slovakia group corrected the form immediately (to *koule*) and one Slovak in the Czechia group managed to say *ko* afterwards, before the recording stopped. We can speculate that maintaining the sound *g* is due to a lower frequency or lower salience of the analogy in these languages. In other words, both languages contain just a few expressions in which the *g* — *k* analogy is present. In contrast, the analogies *u* — *ou* and *a* — *e* affect many more words.

The form *netopýř* was a response to a picture depicting a bat. The Czech word for bat is *netopýř* and its Slovak equivalent is *netopier*. The cognate contains two interlanguage analogies: *ne* [nɛ] — *ne* [ne] at the beginning and *ie* — *ý* before the last consonant. In the non-word *netopýř* both of them appeared in their Czech version. However, the non-word contained also the Czech version of *r* — *ř* analogy which is not present in the target word *netopýř*. Generally, the Slovak language is missing the sound *ř* completely, hence the Czech *ř* sound corresponds always to *r* in Slovak cognates (e.g. Slovak *vejár*, Czech *vějíř* ‚fan‘ or Slovak *halier*, Czech *halíř* ‚penny‘).

#### COMPLETE INTRUSIONS

Our data show that Slovaks in Slovakia produced more intrusions as compared to Slovaks in Czechia (compare 21 Slovaks in Slovakia producing 10 or more intrusions to 8 Slovaks in Czechia with the same pattern). Intrusions of both non-cognates and non-identical cognates can be found in our data. An example of the first type is to say the Slovak word *lienka* instead of the Czech equivalent *beruška* ‚ladybug‘. An example of

the second type is the production of the Slovak word *korok* instead of the Czech form *korek* ‚cork‘. Interestingly, the example of *korok* instead of *korek* demonstrates that intrusions affect also words with interlanguage analogies.

Additionally, a number of responses which are coded as correct data could be potentially treated as intrusions too. It concerns correctly named identical cognates, such as *hruška* ‚pear‘. The production of an identical cognate can be in fact an intrusion, which accidentally corresponds to a correct reaction (expected from a point of view of the researcher). To corroborate whether an identical cognate represents an intrusion requires a deeper analysis. An indication that it might sometimes happen comes from some participants' comments provided in the follow-up interviews. Several participants commented that they were confused by the high number of identical cognates used in the experiment and one of them noted that s/he got the feeling that s/he used only Slovak when naming the pictures requiring identical cognates as a response. Notably, one participant (in Slovakia) admitted that s/he produced several words in Slovak under the tentative assumption that the form could also be a Czech word. Similarly, another participant (in Czechia) said openly that when s/he did not know any Czech word for a picture, s/he produced a Slovak form as a probabilistic guess that it may be the correct answer.

In sum, it seems that identical cognates intrusions can emerge as a deliberate and probabilistic try to arrive at the right answer. The strategy can be summarized as follows: 'I do not know the Czech name for an object, so I try to say a Slovak word in the hope that Czech uses the same form' (see also the discussion in the next section).

#### L1/L2 BLENDS & INTRUSIONS: REPRESENTATION AND SPEECH WORD PRODUCTION

We relate the interpretation of the emergence of L1/L2 blends and intrusions to how a Czech word is represented in the mental lexicon. It can be stated that a Czech word has a low activation level (long-term due to lower proficiency or short-term due to speech situation) or its representation is a) blurred, b) wrong or c) completely missing. Needless to say, the exact interpretation in a particular case is not always straightforward and unequivocal and an output can be explained in different ways.

Furthermore, we postulate that the interlanguage analogies as interiorized mental entities play a part in the creation of a word representation or (also) in the production process. Concerning the word representation, if we agree that the incorrect output *tulen*, for instance (see above), reflects the word representation in the speaker's mind (a wrong representation), it is reasonable to assume that such representation arose by applying the correspondence *-eň* — *-en* (here inappropriately, of course). Regarding the production process, the interlanguage analogies can be involved at its various stages and can serve as a compensational strategy for arriving at specific output when a word representation is blurred, inaccessible or missing. See the following example where this strategy was not successful and where a rule application appears at a later stage, after a form was already produced: a participant produced the (correct) word *sud* ‚barrel‘ (identical cognate) and immediately changed it to *soud* (incorrect form), employing the correspondence *u* — *ou*.





The instances in which *soud* was produced straight away indicate either the wrong representation or the same production process as described above (*sud* > *soud*), with a difference that its first stages took place before the articulation. Under the latter assumption, we conclude that the representation is missing.

The wrong or missing representation can be easily linked not only to the blends of type 1 (which is *soud* and similar cases), but also to errors of type 2 (*studně*), 3 (*dumbír*) and 6 (*goule*) and to complete Slovak intrusions (non-cognates, non-identical cognates, identical cognates).

Intrusions and L1/L2 blends (types 1, 2, 3 and 6) both show indirectly what general strategy Slovak speakers resort to when experiencing difficulties in recalling L2 Czech target. The strategy lies in recalling a Slovak word and a) using it (discussed in the previous section about intrusions) or b) transforming it and using it (L1/L2 blends). The transformation means to turn the Slovak form (or move it closer) to the Czech equivalent by intuitively applying the interlanguage rules. It is worth noting that one participant in Slovakia explicitly asked whether s/he should produce a Slovak word or its transformed form when feeling unsure about a Czech equivalent. Others (Slovaks in Czechia) claimed that they recalled a Slovak form first and immediately translated it into Czech and produced the correct form. Another participant, this time in Czechia, stated that in her/his everyday language use, s/he is sometimes not sure whether s/he uses a Czech word or a Slovak word in some way adjusted to Czech. S/he had similar doubts about certain words even in the experiment.

The speakers are well aware that many words in Czech and Slovak are cognates. This can cause uncertainty relating to what notions are expressed by the same word and an L1 form can be perceived as a safer and more natural option when the participants feel they should produce something.

L1/L2 blends of type 4 (*měruňka*) and 5 (*brusla*) are not part of this strategy as they are based on Czech material. In these cases, it seems that the Czech word representation is either wrong or blurred. One participant (in Slovakia) reported that in several cases s/he knew that the Czech equivalent is very different but did not know its exact form and therefore remained silent. It is conceivable that in other such cases someone may try to (re)construct the form by using various interlanguage correspondences.

The cases of *brusla* and *židla* involve the Slovak part of the interlanguage analogy (in certain sound contexts, the ending *-a* is typical for Slovak words, whereas *-e* occurs in their Czech equivalents). It can be taken as evidence that the representation of the rule was not formed or is not strong enough. However, at the same time, we cannot exclude the possibility that the rule was just applied wrongly or not at all during the language processing. One case of *židla* and one case of *brusla* were immediately corrected to the existing Czech words *židle* and *brusle*, indicating that the correct Czech representation is generally available to the participants and the error relates to processing (related to lower activation level of the representation). Alternatively, we can still assume that the representation is highly blurred and the final output is assembled by using various online strategies and probabilistic guesses. Under such a view, the ending *-a* can be rejected based purely on the reasoning that *-e* seems to be more appropriate for a Czech word.



Notably, self-corrections made immediately after producing non-target responses were observed even for other types of errors, i.e. the Slovak word *orol* was corrected immediately to its Czech equivalent *orel* ‚eagle‘ (intrusion of non-identical cognate), the non-word *slŭn* was corrected to the identical cognate *slon* ‚elephant‘ (L1/L2 blend instead of identical cognate), the non-word *goule* was corrected to the existing Czech word *koule* (L1/L2 blend instead of non-identical cognate; the Slovak equivalent is *gula*) and the production of the non-word *berouš[ka]* was corrected to the existing Czech word *beruška* ‚ladybug‘ (L1/L2 blend instead of non-cognate). Although each example is specific, we argue that the corrections give evidence of a global self-monitoring feedback, comparing and evaluating the output of both inner and outer speech with the production plan (intention) or the stored representations of words (similar to those commented on in Nozari, Dell and Schwartz 2011, see above). The feedback can be applied freely because the task provides sufficient time to arrive at the final decision (5000ms for the production of the name of the picture is a rather long interval).

The existence of such feedback can be used also for the interpretation of the following case (fitting the line about starting with a Slovak form as well). One participant living in Slovakia named the picture of a seal correctly (*tuleň*), which is an identical cognate (see also the description of type 1 blends above), but the same participant reported that after having produced the word, s/he started feeling unsure whether it was correct and subsequently opted for the “I do not know” answer. It is possible that the reason why this participant rejected the initially produced word as a Czech expression is that s/he acknowledged that the analogy *eň* — *en*, which is in many cases correct (i.e. Slovak *kameň* — Czech *kámen* ‚stone‘, Slovak *kmeň* — Czech *kmen* ‚trunk‘, Slovak *remeň* — Czech *řemen* ‚belt‘), was ignored by mistake. At the same time, the participant was not able to recall the correct variant: in searching for the word while applying the analogy, the recall of the form could not be successful because it was not supported by the evidence encountered in the past since the word *tulen* does not exist in either language. As a result, the participant reported not knowing the answer. Under this notion, s/he might have applied the interlanguage analogy post-hoc, on a word already produced.

Self-corrections, especially the corrections of intrusions, can be taken as examples of a lower activation level of L2 word representations. Some participants (from the group in Czechia) mentioned in the follow-up interview that they named a Slovak word because they could not recall a corresponding Czech equivalent within the time limit, even though they know it (and sometimes they produced it in the interview).

In terms of the distinction between the origin of errors at the lexical (word selection) or conceptual (language selection) level, see Zheng, Roelfs and Lemhöfer (2020), the intrusions in our experiment can be attributed to both of them. The lower activation level of L2 words representation discussed in the previous paragraph relates to the former (lexical level). Our data from the follow-up interviews in both participant groups are also in line with the view that part of intrusions originates from the conceptual level. This view is supported by the participants’ comments about being confused by the high number of identical cognates in the experiment (see the section on intrusions). Thus, although the task was focused on the production of L2 words



only, the inclusion of the high number of identical cognates could interfere with the participants' decision-making in language selection for production.

## 6 FINAL DISCUSSION AND CONCLUSIONS

The paper is focused on a qualitative error analysis during a L2 picture-naming task in bilinguals living in the context of receptive bilingualism which involves two closely-related languages. Our main goal was to describe a production error type specific to such a context: systematic L1/L2 blends with frequent interlanguage analogies/rules. Combining L1 and L2 material can be facilitated by the fact the language mixing is not uncommon in the Czech-Slovak context, which can be connected to a lower language cue weight (see De Bot and Schreuder 1993).

At the same time, we find it plausible that speakers derive the rules from the observations of systematicities in the differences between the languages. The locus of the rules in the speech production system is not clear but it is evident that they are employed in various directions when trying to produce L2 words. We postulate that they can be encoded at the lemma level with a particular morphological representation, or separately as a unique representation of language content. They interact with recalling the word's mental representation. We presented several ways of how L2 words can be represented and what forms the interaction between interlanguage analogies and the representations during word production can take. We do not say that one interpretation must be superior to others. In fact, each of them can be plausible as our data contain various kinds of evidence supporting all of them. Hence, the concrete output can be the result of several factors, including linguistic and sociodemographic aspects. Among those, we can list word frequencies, word familiarity or language proficiency.

At first glance, the application of interlanguage analogies corresponds to what Costa, Santesteban and Caño (2005) term lexical-morphological origin of the cognate effect. The authors cite Kirsner, Lalord and Hird (1993), who assume the existence of a shared stem and transformation rules between languages, ensuring the production of only one correct word. Indeed, it is similar to what we assume with regard to Slovak-Czech receptive bilinguals; however, as we saw, the Czech-Slovak interlanguage rules are applied even in non-cognates and in the word stems, which makes the interpretation under this view problematic (how would a stem be shared when it is different across languages?).

Admittedly, the current study is not the first source to comment on the issue of interlanguage analogies. In fact, Sloboda's (2004) theoretical reasoning about the matter was also empirically grounded. However, in relation to Czech and Slovak, our study is the first one to demonstrate the interlanguage analogies at play, using a standard psycholinguistic empirical paradigm. Moreover, we believe that it is the first thorough examination of speech errors of this kind in the literature on psycholinguistic aspects of bilingualism.

However, more research is needed, not only to investigate the Slovak-Czech material more deeply but also other languages sharing cognates which systematically

differ from each other. Various questions follow from our current findings, such as for example:

- (1) Is the frequency of various analogies in a language related to the frequency with which they occur in speech errors?
- (2) Does L2 proficiency play a role? Are speakers with lower L2 proficiency more prone to such errors than more balanced bilinguals?
- (3) Is the probability of an error for a given word related to the speech onset times in the L2 picture-naming task?

We can also generally ask whether the usage of analogies is at play under certain conditions automatically, which would translate into longer reaction times in cases where they are inappropriate. In this respect, we can propose the hypothesis that the production of identical cognates in L2 picture-naming task leads to longer reaction times than the production of non-identical cognates with interlanguage correspondences, at least in Slovaks who are less proficient in Czech and less familiar with Czech words. Such a view is not in accord with claims about the processing advantage of identical cognates over non-identical cognates (e.g., Mulder, Dijkstra and Baayen 2015) or claims about faster production of words with greater formal overlap (Allen and Conklin 2013).

However, we derive our hypothesis from two pieces of evidence. First, we saw that the interlanguage rules can be (wrongly) applied also in identical cognates. In some instances, this pattern allows the interpretation that a Slovak word represents the starting point of the production (which is supported also by several participants in the follow-up interviews). Second, the self-corrections show that the production process involves some sort of feedback (and in fact, the feedback or monitoring is implemented in many models of word production, see Nozari, Dell and Schwartz 2011). Thus, it can be postulated that a speaker chooses a Slovak form and modifies it according to interlanguage correspondences which they frequently encounter in various words. The modification is checked by the feedback procedure and is then accepted (in the case of non-identical cognates) or rejected (in the case of identical cognates, which do not require any modification). The rejection slows down the process and results in longer latencies. This mechanism can be modulated by the language proficiency (Slovaks with lower Czech proficiency could show stronger effects), word frequency/familiarity (less frequent/familiar words would lead to stronger effects) or interlanguage correspondence frequency (words with more frequent correspondences could exhibit stronger effects).

Furthermore, the attention of future research can be focused on similar sociolinguistic contexts, for instance, Danish and Norwegian (see correspondences such as *d/t* in *fod* – *fot*), but also on different situations: is there any acquired rule of the type *x/s*, as in *explore* – *esplorare*, *existence* – *esistenza*, which would manifest itself in the speech of English-Italian bilinguals? The investigation of similar questions could be beneficial for developing a better understanding of bilingual processing, production and representation.

To ensure replicability, requests for the stimuli we used and the data supporting the findings can be addressed to the authors.





## REFERENCES

- Allen, D. B. and K. Conklin (2013) Cross-linguistic similarity and task demands in Japanese-English bilingual processing. *PLoS One* 8/8, e72631.
- Antón-Méndez, I. and T. H. Gollan (2010) Not just semantics: Strong frequency and weak cognate effects on semantic association in bilinguals. *Memory & Cognition* 38/6, 723–739.
- Braunmüller, K. (2002) Semmicommunication and accommodation: observations from the linguistic situation in Scandinavia. *International Journal of Applied Linguistics* 12/1, 1–23.
- Brodeur, M. B., E. Dionne-Dostie, T. Montreuil and M. Lepage (2010) The Bank of Standardized Stimuli (BOSS), a new set of 480 normative photos of objects to be used as visual stimuli in cognitive research. *PLoS One* 5/5, e10773.
- Brodeur, M. B., K. Guérard and M. Bouras (2014) Bank of standardized stimuli (BOSS) phase II: 930 new normative photos. *PLoS One* 9/9, e106953.
- Broersma, M., D. Carter and D. J. Acheson (2016) Cognate Costs in Bilingual Speech Production: Evidence from Language Switching. *Frontiers in Psychology* 7/1461.
- Comesaña, M., A. P. Soares and C. Lima (2010) Semantic representations of new cognate vs. noncognate words: Evidence from two second language learning methods. *Procedia — Social and Behavioral Sciences* 5, 199–203.
- Costa, A. and M. Santesteban (2004) Lexical access in bilingual speech production: Evidence from language switching in highly proficient bilinguals and L2 learners. *Journal of Memory and Language* 50/4, 491–511.
- Costa, A., M. Santesteban and A. Caño (2005) On the facilitatory effects of cognate words in bilingual speech production. *Brain and Language* 94/1, 94–103.
- De Bot, K. (1992) A bilingual production model: Levelt's "speaking" model adapted. *Applied Linguistics* 13/1, 1–24.
- De Bot, K. and R. Schreuder (1993) Word production and the bilingual lexicon. In: Schreuder, R. and B. Weltens (eds) *The bilingual lexicon*, 191–214. Amsterdam/Philadelphia: John Benjamins.
- De Groot, A. M. B. (2011) *Language and Cognition in Bilinguals and Multilinguals*. New York: Psychology Press.
- Dell, G. S. and J. Cholin (2012) Language production. In: Spivey, M. J., K. McRae and M. F. Joannisse (eds) *The Cambridge Handbook of Psycholinguistics*, 426–442. Cambridge: Cambridge University Press.
- Dickins, T. (2009) Češi a slovenština. *Naše společnost* 7/1, 12–26.
- Forster, K. I. and J. C. Forster (2003) DMDX: A Windows display program with millisecond accuracy. *Behavior Research Methods, Instruments, & Computers* 35/1, 116–124.
- Frances, C., E. Navarra-Barindelli and C. D. Martin (2021) Inhibitory and facilitatory effects of phonological and orthographic similarity on L2 word recognition across modalities in bilinguals. *Scientific Reports* 11/12812.
- Gollan, T. H., A. Stasenko, C. Li and D. P. Salmon (2017) Bilingual language intrusions and other speech errors in Alzheimer's disease. *Brain and Cognition* 118/5, 27–44.
- Grasso, S. M., E. D. Peña, L. M. Bedore, J. G. Hixon and Z. Griffin (2018) Cross-linguistic cognate production in Spanish–English bilingual children with and without specific language impairment. *Journal of Speech, Language, and Hearing Research* 61/3, 619–633.
- Haugen, E. (1966) Semmicommunication: the language gap in Scandinavia. *Sociological Inquiry* 36/2, 280–297.
- Hoshino, N., A. L. Beatty-Martínez, C. A. Navarro-Torres and J. F. Kroll (2021) Do Cross-Language Script Differences Enable Bilinguals to Function Selectively When Speaking in One Language Alone? *Frontiers in Communication* 6/668381.
- Hoshino, N. and J. F. Kroll (2008) Cognate effects in picture naming: does cross-language activation survive a change of script? *Cognition* 106/1, 501–511.



- Kirsner, K., E. Lalor and K. Hird (1993) The bilingual lexicon: Exercise, meaning and morphology. In: Schreuder, R. and B. Weltens (eds) *The Bilingual Lexicon*, 215–246. Amsterdam: John Benjamins.
- Klaus, J., K. Lemhöfer and H. Schriefers (2018) The second language interferes with picture naming in the first language: evidence for L2 activation during L1 production. *Language, Cognition and Neuroscience* 33/7, 867–877.
- Kroll, J. F., S. C. Bobb, M. Misra and T. Guo (2008) Language selection in bilingual speech: evidence for inhibitory processes. *Acta Psychologica* 128/3, 416–430.
- Lemhöfer, K., L. Huestegge and K. Mulder (2018) Another cup of TEE? The processing of second language near-cognates in first language reading. *Language, Cognition and Neuroscience* 33/8, 968–991.
- Levelt, W. J. M. (1989) *Speaking: From Intention to Articulation*. Cambridge: The MIT Press.
- Li, C. and T. H. Gollan (2018) Cognates interfere with language selection but enhance monitoring in connected speech. *Memory & Cognition* 46/6, 923–939.
- McMillen, S., Z. Griffin, E. Peña, L. Bedore and G. Oppenheim (2020) “Did I Say Cherry?” Error Patterns on a Blocked Cyclic Naming Task for Bilingual Children With and Without Developmental Language Disorder. *Journal of Speech, Language, and Hearing Research* 63/4, 1–17.
- Mulder, K., T. Dijkstra and R. H. Baayen (2015) Cross-language activation of morphological relatives in cognates: The role of orthographic overlap and task-related processing. *Frontiers in Human Neuroscience* 9/16.
- Musilová, K. (2006) Slovenští vysokoškoláci v ČR (výsledky ankety). In: Nábělková, M. and J. Pátková (eds) *Česko-slovenská současnost a česká slovakistika. Zborník prác k 10. výročiu pražskej univerzitnej slovakistiky*, 102–130. Praha: Filozofická fakulta UK.
- Muyllé, M., E. Van Assche and R. Hartsuiker (2022) Comparing the cognate effect in spoken and written second language word production. *Bilingualism: Language and Cognition* 25/1, 93–107.
- Nábělková, M. (2008) *Slovenčina a čeština v kontakte*. Bratislava/Praha: Veda.
- Nábělková, M. and M. Sloboda (2008) Česko-slovenská komunikácia: semikomunikácia, bivalentnosť a reflexia vzájomnej zrozumiteľnosti slovenčiny a češtiny. In: Tóth, A. J. and A. Uhrinová (eds) *Slovenčina v menšinovom prostredí*, 156–165. Békéscsaba: Výskumný ústav Slovákov v Maďarsku.
- Nozari, N., G. S. Dell and M. F. Schwartz (2011) Is comprehension necessary for error detection? A conflict-based account of monitoring in speech production. *Cognitive psychology* 63/1, 1–33.
- Poarch, G. J. and J. G. van Hell (2012) Cross-language activation in children’s speech production: Evidence from second language learners, bilinguals, and trilinguals. *Journal of Experimental Child Psychology* 111/3, 419–438.
- Poullisse, N. (2000) Slips of the tongue in first and second language production. *Studia Linguistica* 54/2, 136–149.
- Protopapas, A. (2007) CheckVocal: A program to facilitate checking the accuracy and response time of vocal responses from DMDX. *Behavior Research Methods* 39/4, 859–862.
- Roelofs, A. (2018) A unified computational account of cumulative semantic, semantic blocking, and semantic distractor effects in picture naming. *Cognition*, 172, s. 59–72.
- Sloboda, M. (2004) Slovensko-česká (semi)komunikace a vzájemná (ne)rozumiteľnosť. *Čeština doma a ve světě* 12/3–4, 208–220.
- Sloboda, M. (2005) Od aspektov bilingválnej komunikácie k jazykovej asimilácii a retencii: Prípadová štúdia slovenskej rodiny v Česku. *Slovenská reč* 70/6, 338–354.
- Woolard, K. A. (1999) Simultaneity and bivalency as strategies in bilingualism. *Journal of Linguistic Anthropology* 8/1, 3–29.
- Zeevaert, L. and J. D. ten Thije (2007) Introduction. In: Thije, J. D. ten and L. Zeevaert (eds) *Receptive Multilingualism. Linguistic analyses, language policies and didactic*



*concepts*, 1–25. Amsterdam/Philadelphia: John Benjamins.

Zheng, X., A. Roelofs and K. Lemhöfer (2020)

Language selection contributes to intrusion

errors in speaking: Evidence from picture naming. *Bilingualism: Language and Cognition* 23/4, 788–800.

### **Adam Kříž**

Institute of Czech Language and Theory of Communication

Faculty of Arts, Charles University

nám. Jana Palacha 2, Praha 1, Czechia 116 38

ORCID ID: 0000-0002-9349-5342

adam.kriz@ff.cuni.cz

### **Jan Chromý**

Institute of Czech Language and Theory of Communication

Faculty of Arts, Charles University

nám. Jana Palacha 2, Praha 1, Czechia 116 38

ORCID ID: 0000-0001-6646-8026

jan.chromy@ff.cuni.cz