ORIGINAL PAPER

Simulating Code-switching Using a Neural Network Model of Bilingual Sentence Production

Chara Tsoukala[1] · Mirjam Broersma[1] · Antal van den Bosch[2] · Stefan L. Frank[1]

Abstract

Code-switching is the alternation from one language to the other during bilingual speech. We present a novel method of researching this phenomenon using computational cognitive modeling. We trained a neural network of bilingual sentence production to simulate early balanced Spanish–English bilinguals, late speakers of English who have Spanish as a dominant native language, and late speakers of Spanish who have English as a dominant native language. The model produced code-switches even though it was not exposed to code-switched input. The simulations predicted how code-switching patterns differ between early balanced and late non-balanced bilinguals; the balanced bilingual simulation code-switches considerably more frequently, which is in line with what has been observed in human speech production. Additionally, we compared the patterns produced by the simulations with two corpora of spontaneous bilingual speech and identified noticeable commonalities and differences. To our knowledge, this is the first computational cognitive model simulating the code-switched production of non-balanced bilinguals and comparing the simulated production of balanced and non-balanced bilinguals with that of human bilinguals.

Introduction

Bilingual speakers are able to switch from one language to the other, between or within sentences, when conversing with other bilinguals who speak the same languages. This process is called code-switching and it is common among communities where two languages come in contact. For instance, Spanish–English code-switching occurs frequently in the USA among Puerto Rican-Americans (Poplack 1980) and Mexican-Americans (Pfaff 1979), French–Arabic code-switching is common in Morocco (Bentahila 1983) and Algeria (Heath 1984), and Hindi–English in India (Malhotra 1980).

✉ Stefan L. Frank
s.frank@let.ru.nl

Chara Tsoukala
c.tsoukala@let.ru.nl

[1] Centre for Language Studies, Radboud University, Nijmegen, The Netherlands

[2] KNAW Meertens Institute, Amsterdam, The Netherlands

It is incorrect to think of code-switching as a speech error; bilinguals only code-switch when conversing with others who speak the same languages. Grosjean (1997) suggested that bilinguals utilize their languages differently depending on whom they talk to: when they converse with someone with whom they only share one language, they are in a monolingual language mode. When, on the other hand, they are in a setting in which everybody speaks the same languages, they are in a bilingual mode which allows them to code-switch. The amount of code-switching differs per speaker, depending on their personality as well as on the environment and the context of the conversation (Dewaele and Li 2014).

Another misconception is that bilingual speakers mostly code-switch to fill in lexical gaps; this is not the case for proficient speakers (Romaine 1986). However, in the early stages of language acquisition, speakers code-switch more from their less proficient language into their dominant one, rather than vice versa, because they lack the linguistic structures and lexicon needed to communicate; this has been observed both in child bilingual acquisition (Genesee et al. 1995; Petersen 1988) and second language (henceforth L2) acquisition (Sert 2005).

Our goal is to simulate code-switching using a computational cognitive model, with the ultimate aim to further our understanding of the underlying cognitive processes. In a previous study (Tsoukala et al. 2019), we showed that a neural network model designed and validated for monolingual sentence production can generate realistic code-switches when extended to the bilingual case, by training it with syntactic properties and lexical items from two languages and by equipping it with a language control node that allows for the production of either language. Interestingly, the code-switches occurred in the simulations even without exposure to code-switched sentences. In the current study, the aim is to expand on this study in the following three ways: First, we will test the robustness of the model's code-switching behavior; we will do this by replicating the study while randomly varying free parameters. Second, we will simulate balanced and non-balanced bilinguals, and shed light on the code-switching patterns of each simulated group. During the early stages of L2 learning, we hypothesize that the non-balanced bilingual models will code-switch more from their non-proficient L2 into their dominant native language rather than the other way around, possibly because of gaps in their knowledge of the L2. At the later stages of acquisition, however, when the non-balanced models are more proficient in their L2, we hypothesize that these models will also be able to produce code-switches into their L2, even though we expect to find differences in the frequency and distributions of code-switches between the balanced and non-balanced simulation conditions. Third, we will investigate to what extent the simulated code-switches correspond to what is observed in bilingual speech corpora; this is an exploratory analysis with the goal to validate the model patterns.

Code-switching in Balanced and Non-balanced Bilinguals

Code-switching has been studied mainly in early bilinguals, specifically in (i) early balanced bilinguals, i.e., people who have acquired both languages from birth or in early childhood, and (ii) heritage speakers whose home language is a minority language (e.g., Spanish in the USA) and whose dominant language is usually the one spoken in the community (e.g., English) (see, e.g., Poplack 1978; Bullock and Toribio 2009). Speakers who are exposed to an L2 at a later age (e.g., during adulthood) can also code-switch, although the frequency and patterns are known to be different than in early balanced bilinguals. Globalization and greater mobility have caused an increase in the numbers of late non-balanced speakers and there is no social pressure to refrain from code-switching (Matras 2013).

Most studies comparing balanced and non-balanced bilinguals have focused on comprehension rather than production, specifically on reading comprehension (e.g., using

eye-tracking) or grammaticality judgment of code-switched sentences. For instance, Lederberg and Morales (1985) asked different groups of Spanish–English bilinguals to rate the grammaticality of code-switches, and correct them if needed; they compared bilingual children, early balanced bilingual adults, and late non-balanced bilingual speakers who had Spanish as a native language (hereinafter referred to as L1) and moved to the USA as adults where they acquired their L2 English. They found that the late non-balanced bilinguals showed differences in the (grammaticality judgment) acceptance rates compared with the balanced bilinguals; however, the code-switching patterns that the groups followed were similar, which led to the conclusion that the rules governing code-switching are not based on extensive exposure to code-switching, but rather on "knowledge of the grammars of the two code-switched languages in combination with some general linguistic knowledge" (Lederberg and Morales 1985, p.134).

Guzzardo Tamargo and Dussias (2013) studied the reading processing of Spanish-to-English auxiliary phrase switches by balanced and non-balanced bilinguals and found no fundamental differences in the processing patterns between the two groups either, even though the non-balanced bilingual group was slower.

Unlike comprehension studies discussed above, production studies do report differences in the code-switched patterns of early balanced and late non-balanced bilinguals. Poplack (1980) analyzed the code-switching production patterns of Spanish–English bilinguals with varying degrees of proficiency who live in the Puerto Rican community in New York. She observed that balanced bilinguals produced more complex code-switches (e.g., in the middle of the sentence) whereas speakers who were less proficient in their L2 were more likely to switch only for idiomatic expressions, tags (e.g., "you know"), and fillers (e.g., "I mean"). Similarly, Lantto (2012) analyzed the speech patterns of 22 Basque-English bilinguals (10 among them were early balanced bilinguals) and observed clear differences between the early balanced and late non-balanced bilingual groups, with the balanced group producing a wider variety of switch patterns. Psycholinguistic studies (e.g., Gollan and Ferreira 2009, using a picture naming task) have also observed that balanced bilinguals code-switch more frequently.

In the current study, we simulate balanced and non-balanced bilinguals using a sentence production model. We investigate whether the simulations yield differences between the balanced and non-balanced bilingual groups that are similar to those observed in the aforementioned linguistic studies on the production of code-switches in human speech. We thus assess whether the non-balanced bilingual models show a large likelihood to code-switch in the early stages of L2 acquisition from L2 to L1 and whether in later stages of bilingual production the likelihood to

code-switch is higher for the balanced than for the non-balanced models.

The switch directionality of a code-switch (i.e., whether a switch is from the L1 to the L2 or vice versa) can be determined either in an incremental (i.e., linear, as in, e.g., Broersma and De Bot 2006) or a hierarchical manner (using, e.g., one of the most influential grammatical models, the Matrix Language-Frame (MLF) model (Myers-Scotton 1993)). In the following simulations, we have taken a linear approach because we do not want to make assumptions about the way structural relations affect processing.

Bilingual Dual-Path Model

The computational cognitive model we employed for this task is the Bilingual Dual-path model (Tsoukala et al. 2017) which is an extension of the Dual-path model (Chang 2002) of monolingual sentence production. The Dual-path model has successfully modeled a wide range of phenomena over the past years: e.g., structural priming and syntax acquisition in English (Chang et al. 2006; Fitz and Chang 2017) and in German (Chang et al. 2015), cross-linguistic differences in word order preference between English and Japanese (Chang 2009), and input and age of acquisition effects in L2 learning (Janciauskas and Chang 2018).

We chose to base our model on the Dual-path architecture not only because of its success in modeling sentence production but also because it is a learning model (a recurrent neural network, RNN), which therefore allows us to investigate whether code-switched production can emerge from exposure to non-code-switched sentences. Note, however, that a next-word generator, i.e., a simple language model based on an RNN alone or any statistical language model, is very unlikely to produce code-switches without being exposed to code-switched sentences, as the transitional probability between two words in different languages would be zero. The Dual-path, on the other hand, contains a semantic stream and a language control on top of the syntactic stream (the RNN); therefore, it might in theory, and does in practice as our work shows, learn to code-switch even without exposure to code-switched input.

Model Architecture

The Bilingual Dual-path model (Fig. 1) learns to express a given message word-by-word (see "Messages" for examples of messages). The model assumes that there are two paths influencing sentence production: (i) a syntactic path (the lower path in Fig. 1, via the "compress" layers), which is a simple recurrent network (Elman 1990), and (ii) a semantic path (the upper path in Fig. 1), which contains information about the thematic roles (e.g., AGENT, RECIPIENT), the concepts they are connected to, and their realization. The syntactic path learns the syntactic patterns of each language, whereas the semantic path learns to map concepts onto words. Additionally, the model receives information about (i) the event semantics that define when an event takes place (e.g. PAST, PROGRESSIVE) and (ii) the target language, through the corresponding node, which acts as the only language control of the model. Specifically, the target language node simulates the conversational setting in which a speaker is interacting (i.e., one target language in a monolingual setting, both languages in a bilingual setting). All layers use the tanh activation function, except for role and output that use softmax.

The two streams, along with the event semantics and target language, work together to produce grammatically correct sentences that express a specific message.
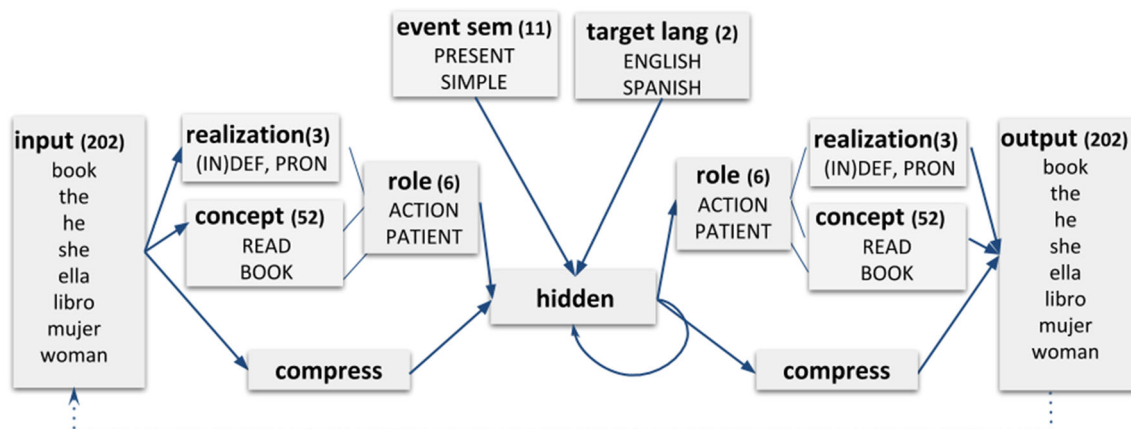
Messages

A message is represented by (i) a target language, (ii) event-semantic information, (iii) pairs of thematic roles and concepts, and (iv) pairs of thematic roles and realizations (pronoun, definite, indefinite) whenever applicable in the case of noun phrases.

The target languages are ENGLISH and SPANISH. The event semantics contain information regarding the aspect (SIMPLE, PROGRESSIVE, PERFECT) and tense (PRESENT, PAST), as well as the thematic roles that are used in each message.

The following simulations make use of 52 unique concepts and six thematic roles: AGENT, PATIENT, AGENT-MODIFIER, RECIPIENT, ACTION-LINKING, and ATTRIBUTE. The roles AGENT and RECIPIENT are only paired with animate nouns (e.g., "son," "cat"). ACTION-LINKING is a combined thematic role that can be used for all main verb types: action (e.g., "shows"), linking ("is"), and possession ("has"). ATTRIBUTE is an attribute expressed only with a linking verb.

Additionally, AGENT, PATIENT, and RECIPIENT are connected to their realization: pronoun (e.g., "he" for the concept BOY), and definite or indefinite article for concepts that are expressed as a noun phrase (e.g., "the boy" or "a boy"). These roles contain optionally a modifier (an adjective, e.g., "a happy dog"). Note that in English the adjective comes before the noun ("the intelligent woman") whereas in Spanish the modifier comes after the noun ("la mujer inteligente"). This knowledge is learned by the model through the training examples and not through explicit syntactic labels.

Fig. 1 The Bilingual Dual-path model receives messages (see "Messages" for examples of messages) and expresses them in sentences, word-by-word. The model is based on the simple recurrent network architecture (the syntactic path, via the "compress" layers), which is augmented with a semantic path that contains information about concepts and their realization, and thematic roles. Additionally, the model receives information on the event semantics and the target language (conversational setting). The numbers in the parentheses indicate the size of each layer (e.g., 52 concept units); the sizes of the hidden and compress layers vary with each training repetition (see "Model Training"). The solid arrows denote trainable connection weights, whereas the lines between roles, realization, and concepts correspond to connections that are given as part of the message-to-be-expressed (e.g., PATIENT is connected to BOOK in a particular message). The dotted arrow indicates that the produced word is given back as input, influencing the production of the next word

(Message-to-)Sentence Production

For a message to be expressed, the following nodes need to be activated in the model: the event semantics (e.g., PRESENT, PAST) and the target language (ENGLISH, SPANISH) that specifies the intended output language. Furthermore, the semantic roles (e.g., ACTION, PATIENT) are connected to their respective concepts (e.g., READ, BOOK) and realizations (e.g., INDEF for an indefinite article). For example, if the message is:

AGENT = WOMAN, DEF
ACTION = GIVE
PATIENT = BOOK, INDEF
RECIPIENT = FATHER, INDEF
TARGET-LANG = ENGLISH
EVENT-SEM = PRESENT, PROGRESSIVE,
            AGENT, PATIENT, RECIPIENT

the model would learn to express it in English as "the woman is giving a book to a father," and if the target language was Spanish as "la mujer está dando un libro a un padre." Following Chang et al. (2006), to express the recipient before the patient ("the woman is giving a father a book" or "la mujer está dando a un padre un libro"), the PATIENT receives less activation through the event semantics, thus prioritizing the RECIPIENT. In the current example, the event semantics would be PRESENT, PROGRESSIVE, AGENT, PATIENT:0.5, RECIPIENT.

When the model receives a message, it produces it word-by-word. The produced word is the output word with the highest activation. Each produced word is then given as input in the next time step, and it influences the production of the next word. The period (".") works as an end-of-sentence marker and the model stops producing words when it outputs the period or if it has exceeded the length of the target sentence, plus 2 extra words. We allow extra words because the model might produce a different structure than the target one; for instance, the message of the sentence "the boy is giving the girl a key" (double dative) could also be expressed as "the boy is giving the key to the girl" (prepositional dative).

Miniature Languages

In order to simulate Spanish–English sentence production, we generated training sentences that are derived from a small subset of the syntactic properties ("Tense and Aspect") and the lexica ("Bilingual Lexicon") of the two languages. Note that the model does not contain a phonological level because we are only focusing on the interaction between semantics and syntax, and not on restrictions imposed by phonology.

Tense and Aspect

The allowed tenses used in the structures are past and present, and the aspects simple, progressive, and perfect. The past tense is only used in simple aspect sentences (e.g., "the girl jumped"), whereas the present tense applies to all three aspects. The allowed structures for the two languages and all tenses and aspects can be found in Table 1.

Table 1 Allowed structures for English and Spanish

Aspect and tense	English example	Spanish example
Simple past		
SV	the brother sneezed	el hermano estornudó
SVO	a gentleman pushed the chair	un señor empujó la silla
Simple present		
SV	the girl swims	la niña nada
SVO	the grandmother kicks the toy	la abuela patea el juguete
SVO (linking)	the hostess is happy	la anfitriona está feliz
SVO (possession)	a dog has the ball	un perro tiene la pelota
SVDI	the man throws a book to the aunt	el hombre tira un libro a la tía
SVID	the man throws the aunt a book	el hombre tira a la tía un libro
Perfect present		
SV	he has walked	él ha caminado
SVO	a woman has pushed the pen	una mujer ha empujado el bolígrafo
SVDI	a girl has thrown a key to the lady	una niña ha tirado una llave a la señora
SVID	a girl has thrown the lady a key	una niña ha tirado a la señora una llave
Progressive perfect		
SV	she is jumping	ella está saltando
SVO	he is kicking a chair	él está pateando una silla
SVDI	the father is giving the toy to a girl	el padre está dando el juguete a una niña
SVID	the father is giving a girl the toy	el padre está dando a una niña el juguete

S, subject; *V*, verb; *O*, object; *D*, direct object; *I*, indirect object

Bilingual Lexicon

The bilingual lexicon (Table 2) is an extension of the lexicon used in the Tsoukala et al. (2019) study. It contains 202 words: 92 English words, 109 Spanish words, and the shared end-of-sentence marker ("."). The Spanish lexicon is larger because Spanish is a gendered language. For instance, nouns and adjectives are usually expressed differently depending on whether they modify a masculine noun or a feminine one (e.g., "busy" is "ocupado" if it modifies a feminine noun and "ocupada" for a masculine noun). We also included four common-gendered Spanish adjectives such as "feliz" ("happy") that do not change depending on the noun it modifies.

The verbs are either intransitive (e.g., "swims"), transitive ("carries"), double ("throws"), linking ("is," "está"), or possession verb ("has," "tiene"). The two linking verbs ("is," "está") and the English possession verb ("has") were also used as auxiliary verbs for the progressive and perfect forms, respectively, as was the Spanish perfect-form auxiliary verb "haber." Following the allowed structures, each verb had four forms: simple present, simple past, present participle, and past participle.

Note that syntactic category information (such as "noun," "participle") is not given explicitly; the model learns through training (via the syntactic path) that words that occur in similar context tend to be of the same syntactic category.

Message-Sentence Pair Examples

We hereby illustrate how a message corresponds to (and is expressed with) a sentence. For instance, the following message:

    AGENT = WAITER, DEF
    AGENT-MOD = TALL
    ACTION-LINKING = SNEEZE
    EVENT-SEM = SIMPLE, PAST, AGENT,
                AGENT-MOD

corresponds to the following sentences in English and Spanish:

– the tall waiter sneezed.
– el camarero alto estornudó. (literally: "the waiter tall sneezed.")

Changing the tense of the message to PRESENT instead of PAST would correspond to the sentences "the tall waiter sneezes" and "el camarero alto estornuda," whereas further changing the aspect to PROGRESSIVE instead of SIMPLE would correspond to "the tall waiter is sneezing" and "el camarero alto está estornudando."

Table 2 Syntactic categories in the bilingual lexicon (Spanish in italics)

Syntactic category	n	Examples
Verbs	66	
Auxiliary	4	is, has, *está, ha*
Intransitive	16	walked, swims, *nada*
Transitive	12	carries, pushed, *lleva*
Double	12	throws, gives, *tira*
Possession	4	has, had, *tiene, tenía*
Linking[a]	2	is, *está*
Participles[b]	56	
Progressive	28	walking, *caminando*
Perfect	28	eaten, *comido*
Nouns	52	
Animate	40	uncle, aunt, *tío, tía*
Inanimate	12	pen, book, *libro*
Prepositions	2	to, *a*[c]
(Predicate) adjectives	26	busy, *ocupado*
Determiners	6	a, the, *un, una, el, la*
Pronouns	4	he, she, *él, ella*

[a]Both linking verbs overlap with the auxiliary verbs

[b]Nine of these have the same form as a verb; e.g., "walked" is either a perfect participle or a verb

[c]The Spanish preposition "*a*" and the English indefinite article "a" are differentiated, in that they have separate nodes in the lexicon

Messages that contain direct and indirect objects can be expressed with the thematic roles of PATIENT and RECIPIENT respectively. For instance, the following message:

> AGENT = FATHER, PRON
> ACTION-LINKING = THROW
> PATIENT = BALL, DEF
> RECIPIENT = DOG, INDEF
> EVENT-SEM = SIMPLE, PRESENT, AGENT,
>   PATIENT, RECIPIENT

is expressed as "he throws the ball to a dog" or "él tira la pelota a un perro."

Finally, messages that contain linking verbs are encoded using an attribute:

> AGENT = MAN, DEF
> AGENT-MOD = KIND
> ACTION-LINKING = BE
> ATTRIBUTE = TIRED
> EVENT-SEM = SIMPLE, PRESENT, AGENT,
>   AGENT-MOD, ATTRIBUTE

which is expressed as "the kind man is tired" or "el hombre amable está cansado."

Model Training

The model learns through supervised training. A message is given as input and the network tries to generate a sentence word-by-word; after a word has been produced, it is compared with the target word and the weights are adjusted according to the backpropagation algorithm. All networks were trained for 40 epochs using 2000 message-sentence pairs.

The backpropagation parameters were the same across all simulations: the momentum was set to 0.9 and the initial learning rate was 0.10, which linearly decreased for 10 epochs until it reached 0.02, at which point it was held constant from epoch 11 onward. This applies to both the Balanced and the Non-balanced bilingual models ("Balanced Bilingual Model (Balanced Model)" and "Non-balanced Bilingual Models (L1 English and L1 Spanish Models)", respectively). Note that the Non-balanced models are exposed to their L2 around the 15th epoch; therefore, they start learning the L2 with a decreased learning rate (0.02).

To increase the generalizability of the reported results and to test the robustness of the results reported in Tsoukala et al. (2019), we trained 40 networks per simulation, randomizing all free parameters (as seen below), excluding the backpropagation parameters (i.e., the momentum and learning rate) and the training set size. The parameters were randomized per training repetition (i.e., for each of the 40 networks), but the same parameter values were kept across the three different simulations: e.g., the first training repetition of the balanced bilingual simulation had the same initialized weights as the first training repetition of the non-balanced bilingual model(s).

First, the message-sentence pairs were randomly generated for each simulation before the training started. The sentences were constrained by a set of allowed structures ("Tense and Aspect") and for each syntactic category a randomly selected word was sampled from the bilingual lexicon ("Bilingual Lexicon"). Note that the target sentences were never code-switched.

Second, when sampling the bilingual training set for the balanced and non-balanced bilingual simulations, we varied the percentage of English and Spanish. The percentage of English was sampled from a normal distribution with a mean of 50% (standard deviation: 8) and the rest was Spanish. Third, weights of trainable connections were initialized using Xavier initialization (Glorot and Bengio 2010). Last, the weights of the connections between thematic roles and concepts ("concept"–"role" and "predicted role"–"predicted concept" in Fig. 1), which are not trained, were integer values sampled between 10 and 20 (the exact value was randomized once per training repetition and was the same for all these connections).

The hidden layer size was also sampled per training repetition (between 70 and 90 units) and the compress layer size was set to the closest integer to 77% of the hidden layer size.

Code-switching

As mentioned above, the target sentences did not contain any code-switches. To allow the model to code-switch, we manipulated the model's language control (target language node) when testing, which simulates the conversational setting, or language mode (Grosjean 1997). Only one of the target languages was activated before the production of the first word, and the network was thereby biased towards producing the first word in that language, but once the first word was produced, both languages were activated. This allowed the model to continue in the same language or to code-switch.

With regard to the code-switching types, in the current simulations, we look at two types of code-switches, which Muysken (2000) calls (lexical) insertions and (intra-sentential) alternations:[1]

– Insertional switching (insertion of single words), e.g., "He gave the *libro* to my niece." (He gave the book to my niece.)
– Alternational switching (intra-sentential switching), e.g., "*María prefiere hacer el viaje* by train instead." (Maria prefers to make the trip by train instead.)

Correctness of Sentence Production

A produced sentence is considered grammatically correct if it consists of an allowed sequence of syntactic categories, i.e., if the sequence exists in the training set. The criterion for correct meaning is that the sentence is grammatical and that all thematic roles are expressed correctly, even if they are code-switched, but with no omitted or extra attributes (e.g., not "dog" instead of "big dog" or vice versa). In some cases, the meaning can be correct even if the produced sentence is different than the target. For instance, if a double dative sentence ("the woman gives the cat a ball") is expressed as a prepositional dative ("the woman gives a ball to the cat"), the meaning is counted as correct because the message is expressed correctly.

---

[1]Muysken also identified other types of code-switches (i.e., congruent lexicalization) and sub-categories of the insertions and alternations (e.g., insertions of fixed expressions, idioms and tags, and alternations between sentences called "inter-sentential switching"), but these fall beyond the scope of the model because the model produces single sentences without context and without the usage of fixed expressions and tags.

Method: Simulations and Corpus Analysis

We addressed the three goals of this study by running three sets of simulations. First, having expanded the lexicon and having varied almost all free parameters compared with the Tsoukala et al. (2019) study (see "Model Training"), we ran 40 training repetitions, with different parameters each, to investigate (i) whether the model again produces code-switched sentences and (ii) the sensitivity of this behavior to the random parameter settings and initial weights.

Second, we simulated balanced and non-balanced Spanish–English bilinguals and compared their production patterns with respect to code-switching. Specifically, we measured (i) how often a sentence is code-switched in total and per switch direction[2] (Spanish-to-English versus English-to-Spanish), (ii) what kind of code-switches (alternational, insertional) are produced and at which syntactic point, and (iii) how the patterns vary with the amount of training and exposure to the two languages. Note that each epoch corresponds to the amount (time) of learning, not the amount of training examples per language; the non-balanced bilingual models are initially exposed only to their L1, whereas the balanced bilingual model directly receives bilingual message-sentence pairs, thus receiving approximately half the exposure per epoch to each individual language. Third, we test the validity of the simulated patterns by comparing them with human data, i.e., code-switched utterances in bilingual speech corpora.

To address the first two goals, we run one early balanced bilingual model and two late non-balanced bilingual models with different L1 (English, Spanish).

For the third goal, we analyzed the Bangor Miami corpus (Deuchar et al. 2014) to obtain code-switched patterns of Spanish–English bilingual speech.

Balanced Bilingual Model (Balanced Model)

The Balanced model was simultaneously exposed to both languages (roughly 50% per language as described in "Model Training"), therefore simulating balanced bilinguals. The Balanced model was trained for 40 epochs using 2000 message-sentence pairs and tested on 500 messages. The training and test sets were unique per training repetition (40 training repetitions in total) and the distribution of Spanish and English in the test messages was the same as in the training messages.

---

[2]As mentioned in "Code-switching in Balanced and Non-balanced Bilinguals," our approach is linear. We start from the first word of a code-switched output sentence; if the word is in Spanish, we mark the switch direction as "Spanish-to-English," whereas if the first word is English we count it as "English-to-Spanish."

Non-balanced Bilingual Models (L1 English and L1 Spanish Models)

The non-balanced bilingual models were first exposed only to their L1 for roughly 15 epochs. Specifically, the L1 English model was trained with English-only sentences (2000 message-sentence pairs) for about 15 epochs (the exact number of epochs was randomly sampled between 13 and 17), whereas the L1 Spanish model was initially trained on Spanish-only sentences (2000 message-sentence pairs). For the remaining epochs (making a total of 40), the networks were exposed to the same 2000 message-sentence pairs as the Balanced model and tested on the same 500 messages. Once again, there were 40 training repetitions per model and the message-sentence pairs and test messages were different for each run.

Corpus Analysis

To compare the simulated patterns to human data, we analyzed the transcriptions of the Bangor Miami corpus (Deuchar et al. 2014)[3] that consists of 56 spontaneous and informal conversations between two-to-five speakers, living in Miami, FL. Out of the 84 speakers, 60 were equally fluent in English and Spanish.[4] Each word in the conversation file has been automatically tagged with a language code (English or Spanish) and a syntactic category (e.g., noun). We selected the sentences that contained more than one language code, resulting in 2796 code-switched sentences, which is 6.2% of the corpus (45,289 sentences in total). We then divided the code-switches into alternations, in case the sentence continued in the code-switched language, and insertions, if the code-switches were single words that were inserted (once or several times) in the sentence. Meanwhile, we corrected the syntactic categories of erroneous or missing tags.[5]

From the 2796 code-switches observed in the Miami corpus, we included only the 1369 that occur at syntactic categories that are relevant to our model; for instance, we excluded interjection insertions because interjections are out of the scope of the model.

As an additional corpus, we compared the model's patterns with the code-switches observed in Poplack (1980). Note that Poplack's corpus is not publicly available; therefore, we could not re-analyze the data. Likewise, out of the 1835 code-switching instances observed in Poplack's

study, we have only included the syntactic categories that are relevant for our study.

Note that in both corpora most switches are so-called extra-sentential, which are not grammatically or semantically related to any other part of the sentence (e.g., tag insertions, such as "you know" and "right?") and are therefore not included in the model.

Results

Model Performance

Balanced Model

Figure 2 shows the performance (i.e., percentage of sentences with correct grammar and with correct grammar and meaning) of the balanced bilingual model on its two native languages: Spanish (Fig. 2a) and English (Fig. 2b). Both languages are learned equally well: the mean percentage of sentences that are produced with correct meaning at the last epoch (hereinafter: correct meaning) is 83% for Spanish and 85.4% for English.

L1 English and L1 Spanish Models

Figure 3 (upper row) shows the performance of the L1 Spanish (Fig. 3a) and L1 English (Fig. 3b) models on their native language. Note that around the 15th epoch the L2 is introduced which slightly affects the production of the L1.

The lower row of Fig. 3 indicates the performance of the non-balanced bilingual models on their L2, starting from the epoch in which the L2 was introduced. Note that because the exact starting epoch varies per training repetition, only after the 18th epoch are all 40 training repetitions introduced to the L2; until then, the plot shows the mean only of the training repetitions that have already been exposed to the L2. Figure 3c shows the performance of L2 English in the L1 Spanish model and Fig. 3d shows the L2 Spanish performance of the L1 English model.

Research Goal 1: Code-switching in the Models

In the final epoch, the Balanced model (Fig. 2) produces 21.4% Spanish-to-English and 27.0% English-to-Spanish code-switches, out of all correctly produced sentences. Examples of code-switched sentences include:[6]

– a boy pushed *la silla* ("the chair")
– a happy cat *tiene una pelota* ("has a ball")
– the uncle *está triste* ("is sad")
– a dog *corrió* ("ran")

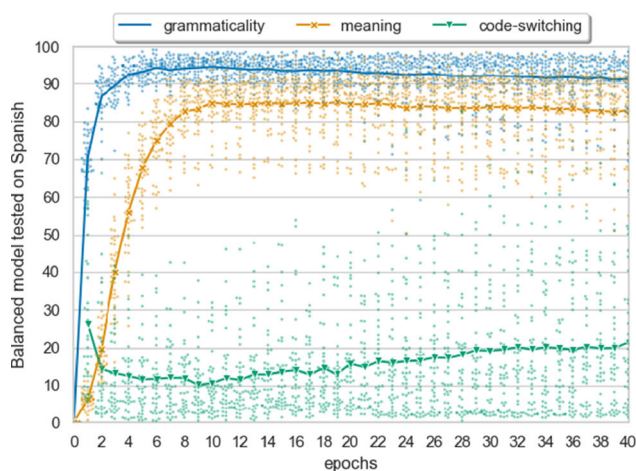---

[3] http://bangortalk.org.uk/speakers.php?c=miami

[4] Fluency was measured by self-reported "Spanish ability" and "English ability." The questionnaire results can be found on the corpus website.
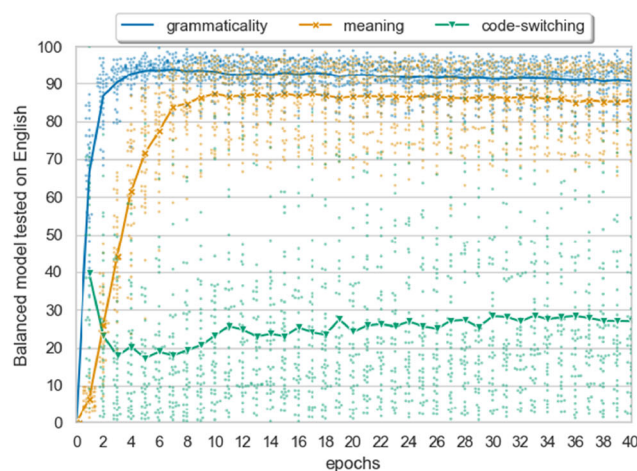
[5] The scripts used, as well as the resulting sentences, can be found at https://osf.io/vd3wa/.

[6] The full list of output sentences per model can be viewed online at https://osf.io/vd3wa/.

(a) Balanced model tested on Spanish

(b) Balanced model tested on English

Fig. 2 Mean grammaticality, correctness of meaning, and code-switch percentage of the balanced bilingual model, tested on Spanish (**a**) and English (**b**). The dots are jittered and represent each individual training repetition

The non-balanced models' code-switching patterns develop over time: at the early stages of L2 learning they produce very few L2 sentences correctly, most of which contain code-switches into the L1; for instance, on the 14th epoch the L1 English model produces 3.5% of Spanish sentences correctly, out of which 87.9% contain a code-switch into English. Respectively, the L1 Spanish model produces 5.9% of English sentences correctly, with 90.5% of these containing a switch into Spanish. Over time, though, the models become more proficient in their L2 and stop reverting to their L1: at the end of the training the L1 Spanish model reaches 55.1% in meaning accuracy of English sentences and produces 5.3% switches into Spanish, whereas the L1 English model reaches 55.9% accuracy on Spanish and switches back into English 3.9% of the time. Code-switches from the L1 into the L2 are more steady throughout acquisition: the L1 English model code-switches 0.9% from English into Spanish, and the L1 Spanish model code-switches 1.3% of the time from Spanish into English.

Research Goal 2: Balanced Versus Non-balanced Model Comparisons

The second goal of this study is to investigate the differences in code-switch types produced by the balanced and non-balanced models at the late stages of acquisition, when both models have been exposed to the bilingual input for 25 epochs (i.e., the 25th epoch for the balanced model versus the 40th epoch for the non-balanced models). Table 3 presents the percentage of the total code-switch types (alternational, insertional, and final-word, in case the switch is at the end of the sentence and it is therefore unclear

whether it is an insertion or an alternation) for the three models (balanced Spanish–English bilingual, non-balanced bilingual with L1 English, non-balanced bilingual with L1 Spanish). The balanced bilingual model code-switches much more frequently than the L1 Spanish and L1 English models.

Figure 4 compares the three models with respect to the switch type and switch direction. The percentages shown here are against all correctly produced sentences *of that target language*, not of all sentences as in Table 3.

Additional information on the exact code-switching patterns per switch type (alternational, insertional, final-word), language direction (English-to-Spanish and Spanish-to-English), and syntactic category in which the switch occurs can be found at https://osf.io/vd3wa/ under results/supplementary_plots.

Research Goal 3: Model Versus Corpus Comparison

To test the validity of the produced patterns, we compared the simulated code-switched patterns with the ones observed in the Miami corpus, as well as in the patterns observed in Poplack's (1980) study. The results can be found in Table 4.

Both corpora contain code-switches from all participants, both balanced bilinguals and Spanish-dominant. To compare the model results with the corpora that contain both balanced and Spanish-dominant speakers, Table 4 reports switches from the Balanced and L1 Spanish models combined. The simulations produce a high percentage of noun phrase alternations, which is also the case in Poplack's study and the Miami corpus. Furthermore, both the corpora and the model display a substantial (but small) amount of verb alternations. The
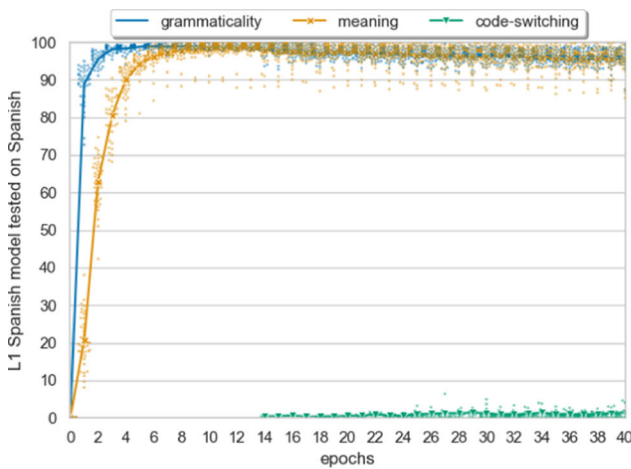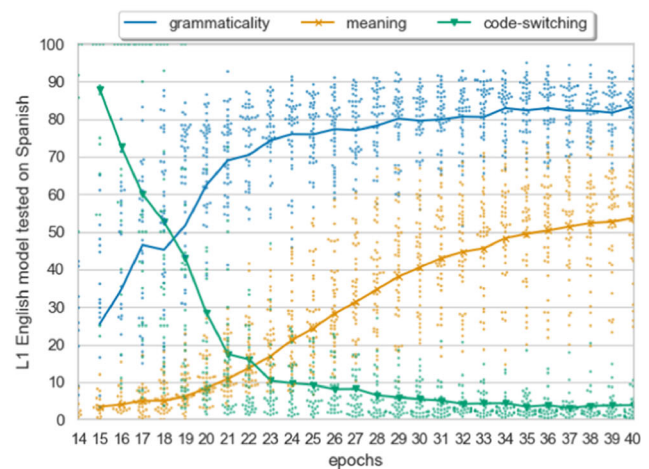
(a) L1 Spanish tested on Spanish
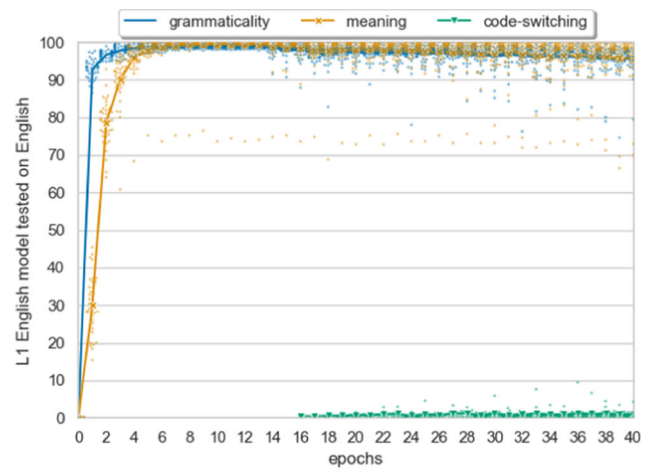
(b) L1 English tested on English
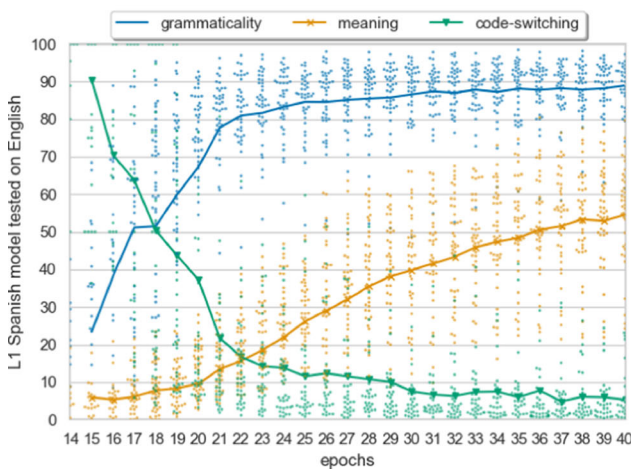
(c) L1 Spanish tested on L2 English

(d) L1 English tested on L2 Spanish

Fig. 3 Mean grammaticality, correctness of meaning, and code-switch percentage of the non-balanced bilingual models, tested on their L1 (L1 Spanish (**a**) and L1 English (**b**)) and L2 languages (L2 English (**c**) and L2 Spanish (**d**)) over 40 training repetitions (L1 Spanish model: left column, L1 English model: right column). The dots are jittered and represent each individual training repetition

other phrase alternations are fewer in the models than in Poplack's data, which is probably due to the fact that with the current artificial languages we have only simulated prepositional phrases whereas the phrase alternation in the corpus include other types of phrases as well (i.e., adjective, adverb and infinitive phrases). Both the

simulations and human bilinguals, especially in Poplack's study, disprefer preposition insertions.

There are also clear differences between simulated and empirical code-switches. Unlike human bilinguals, the model seems to favor determiner insertion, and more specifically Spanish determiners. The most striking difference

Table 3 Percentage of switch types per model over all correctly produced sentences after 25 epochs of exposure to both languages (25th epoch for the Balanced model and 40th for the L1 English and L1 Spanish models)

Switch type	Balanced	L1 English	L1 Spanish
Alternation	15.0% [12.7, 17.9]	1.7% [1.0, 3.1]	1.1% [0.8, 1.6]
Insertion	2.3% [1.8, 2.9]	0.1% [0.1, 0.2]	0.0% [0.0, 0.1]
Final-word	3.7% [2.9, 5.0]	0.3% [0.2, 0.5]	0.2% [0.1, 0.3]
Total	21.0% [17.9, 24.9]	2.1% [1.4, 3.6]	1.3% [1.0, 1.8]

The numbers in the brackets show the 10,000-sample bootstrapped 95% confidence interval

(a) Spanish → English
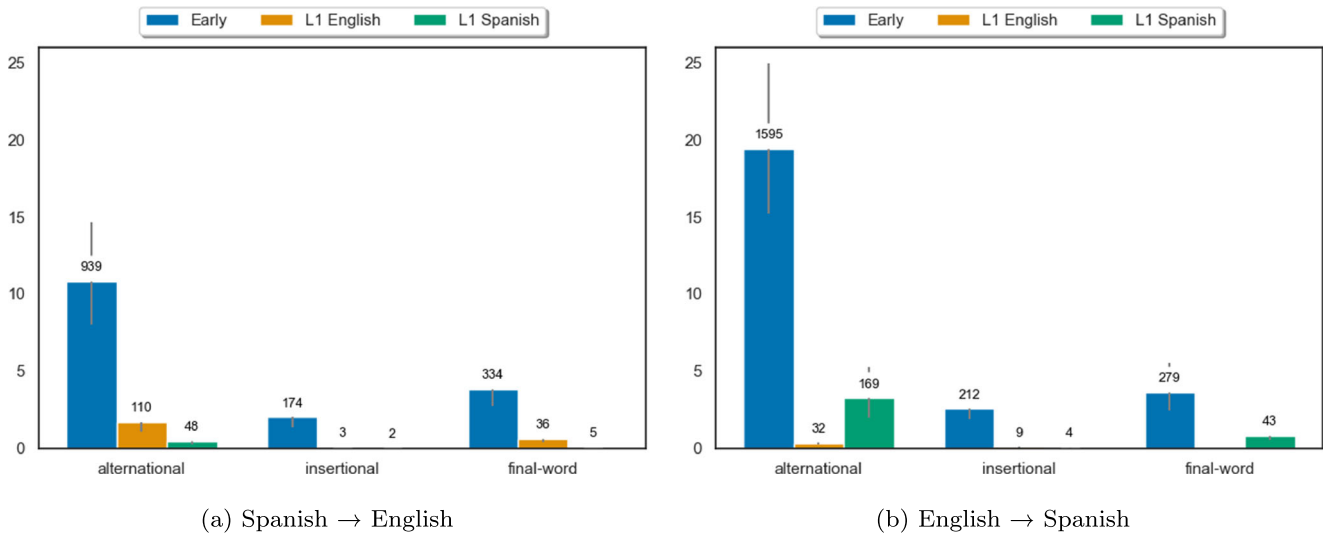
(b) English → Spanish

Fig. 4 Model comparison of code-switching types. The percentage is against all correctly produced sentences of that target language ” (Spanish **(a)** and English **(b)**) and of that model, and the numbers denote the absolute number of sentences with that switch type. The error bars show the 10,000-sample bootstrapped 95% confidence interval

between corpora and model results, however, is found in noun insertions: both corpora showcase that noun insertion is a major switching category among human bilinguals.

The model, however, produces less than 1% of noun insertions, whereas the two corpus data report over 30% of noun insertions.

Table 4 Number of code-switches by syntactic category and language in (i) the balanced bilingual and L1 Spanish models ("Sim."), (ii) the Miami corpus ("Miami"), and (iii) Table 2 of Poplack (1980) ("Pop."), adapted to include only syntactic categories the model produces

Syntactic category of CS	# English to Spanish			# Spanish to English			# Total CS			% of total CS		
	Sim.	Miami	Pop.	Sim.	Miami	Pop.	Sim.	Miami	Pop.	Sim.	Miami	Pop.
Insertions:												
Determiner	102	7	3	49	3	0	151	10	3	4.8	0.7	0.5
Noun	21	120	34	6	367	141	27	487	175	0.9	35.6	31.4
Auxiliary	32	0	0	51	0	0	83	0	0	2.6	0.0	0.0
Verb	44	25	6	58	9	13	102	34	19	3.2	2.5	3.4
Adjective	0	24	3	2	58	12	2	82	15	0.1	6.0	2.7
Preposition	15	9	2	2	10	0	17	19	2	0.5	1.4	0.4
Alternations:												
Noun phrase			106			103			209			37.5
Determiner	954	77		517	51		1471	128		47.0	9.4	
Noun	38	25		8	178		46	203		1.5	14.8	
Adjective	0	30		9	87		9	117		0.3	8.5	
Verb phrase			27			13			40			7.2
Verb	297	104		183	57		480	161		15.3	11.8	
Auxiliary	304			241			545			17.4		
Participle	71			28			99			3.2		
Prepositional phrase	100	68	55	1	60	39	101	128	94	3.2	9.3	16.9
Total	1978	489	236	1155	880	321	3133	1369	557	100.0	100.0	100.0

In alternational switching, the code-switching starts at the syntactic category presented in the leftmost column and continues in the non-target language. For instance, an adjective alternation within a noun phrase means that the adjective of the noun phrase was the first code-switched item of that sentence

Discussion

Code-switching in the Models

The first goal of this paper was to verify the robustness of the code-switch model presented in Tsoukala et al. (2019). Having varied almost all free parameters in the current simulations, and using an expanded lexicon, we tested again whether the bilingual model is able to produce code-switches that are attested in bilingual speech, even without having been exposed to code-switched input. The models indeed produced code-switches, thus confirming that code-switching can partially be explained by the distribution of the two languages involved (in combination with the cognitive architecture of the model, in our simulations). This is in line with Lederberg and Morales (1985), who claimed that (extensive) exposure to code-switching is not needed for a bilingual speaker to code-switch.

Importantly, the model is able to code-switch by merely having (and manipulating) a language control ("target language") node that sets the conversational setting and allows the model to produce in either language. No other cognitive control was required for the model to code-switch.

It is interesting to observe the huge variance in the amount of code-switching between training repetitions; at the last epoch of the balanced bilingual model tested on English (Fig. 2b), the percentage of code-switches produced by the 40 models ranges from 2.3 to 80.8%. Large individual variance is something that has also been observed among human bilinguals (Dewaele and Li 2014).

As mentioned in "Bilingual Dual-Path Model," an RNN alone trained on non-code-switched data is unlikely to produce code-switched sentences. As a case in point, we trained the SRN-only part of the model (i.e., the syntactic stream alone) using the same input and settings as described in the "Method: Simulations and Corpus Analysis" section. It is difficult to directly compare the Dual-path with an SRN-only version because the former expresses a specific message; for an approximate comparison, we gave the SRN-only model the first word of the target message and let it produce any sentence. The SRN-only model learned to produce grammatical sentences but it did not produce any code-switched sentences.

Balanced Versus Non-balanced Model Comparisons

The second aim of this study was two-fold: First, to investigate the development of code-switches over time in the non-balanced bilingual models. Second, to compare the production patterns of balanced and non-balanced bilinguals and per switch direction. On the one hand, at the early stages of L2 acquisition, the non-balanced models have not been exposed enough to their L2 and they strongly prefer

to switch back into their L1 (i.e., over 87% of the time). This preference is in line with what has been observed in bilingual language acquisition by children (e.g., Petersen 1988). When comparing, on the other hand, the balanced and non-balanced models after an equal amount of exposure with both languages (25 epochs), the patterns change: the balanced bilingual model code-switches considerably more frequently than the non-balanced bilingual models, which is in line with what has been observed in humans (e.g., Poplack, 1980, Gollan and Ferreira 2009).

Note that the non-balanced bilingual models perform better in their L1 compared with the balanced bilingual model: 95.5% accuracy in meaning in the last epoch for the L1 Spanish model and 95.9% for the L1 English model on their L1 (Fig. 3), as opposed to 85.4% for English and 83% for Spanish accuracy in the Balanced model (Fig. 2). The reason behind this discrepancy is that the non-balanced bilingual models receive double the input in their L1 (for the first 15 epochs) compared with the balanced bilingual model that has two native languages. As mentioned above, an epoch corresponds to the amount of learning time, not the input received.

In the current simulations, we have assumed that the L1 is the dominant language. However, a large proportion of bilingual speakers in communities that code-switch are heritage speakers who, as mentioned in "Code-switching in Balanced and Non-balanced Bilinguals," are more exposed to (and fluent in) their L2, the majority language of the country they live in, rather than the L1 that is mostly spoken at home. Heritage speakers could also be simulated in the model, by first exposing the model to the L1 only (similar to the non-balanced models) and then introducing bilingual input in which the L2 is much more frequent than the L1, reflecting heritage speakers' exposure.

Model Versus Corpus Comparison

The third goal of this paper was explorative, aiming to validate the model by investigating to what extent the simulated patterns correspond to bilingual speakers' behavior. We cannot expect a perfect match between the model's code-switching patterns and the corpus data because the simulations use an artificial micro version of English and Spanish. Nevertheless, some patterns are similar to what human bilinguals produced in the two corpora (see Table 4 for percentages of patterns of all three studies). There are also noticeable differences between the modeled and human code-switched patterns, with the most striking one being the high amount of noun insertions in the two corpora compared with the simulations. A possible explanation for this discrepancy is that human bilinguals tend to prefer a specific language depending on the domain, for instance, Spanish for food, English for school- and

work-related terms (Fishman et al. 1971). Additionally, bilinguals align with their collocutors and repeat syntactic structures and utterances (Fricke and Kootstra 2016). In the Miami corpus, for instance, when we analyzed the noun insertions per dialogue (chat transcription), we found 117 repetitions out of the 487 insertions. The model, on the other hand, simulates individual sentences and has no context of what has been produced before nor a notion of domain-specific terms; the only context given is the language control, which specifies whether the setting is monolingual or bilingual. Another minor reason behind the small number of noun insertions in the model simulations is that we have excluded from our analysis final-word switches, as we are unable to determine whether they are insertions or alternations. In the corpus analysis, on the other hand, we counted 296 final-word noun switches as noun insertions.

The Role of Community Norms

One advantage of computational modeling as a method of studying code-switching is that we can investigate which patterns are caused by the languages' statistical properties in combination with the cognitive system, rather than by community norms. Community norms have been shown to influence code-switching patterns, to the extent that there can be opposite preferences between communities that use the same language pair; for instance, Blokzijl et al. (2017) analyzed the Miami corpus and a Nicaraguan Spanish–English creole corpus and found that speakers in the Miami corpus preferred to use a Spanish determiner in a mixed determiner phrase, whereas in the Nicaraguan corpus only the English creole determiner was used. Similarly, Balam et al. (2020) examined three Spanish–English communities (from Northern Belize, New Mexico, and Puerto Rico) with respect to their preference for two code-switched compound verb constructions; the US bilinguals showed a different preference to the Northern Belize community, once again confirming that there are aspects other than grammar that affect code-switching across communities that speak the same language pair. Modeling can help disentangle the linguistic from the extra-linguistic factors. For example, our model produced certain patterns and in certain switch directions (e.g., more Spanish determiners than English ones) even though there were no extra-linguistic influences available to the model.

Conclusion

We have shown that the Tsoukala et al. (2019) results are robust: the Bilingual Dual-Path model can produce code-switching patterns without exposure to such code-switched patterns, by only having a language control that allows the model to produce in either language. Furthermore, we simulated three groups of bilinguals and showed the differences between the early balanced and late non-balanced simulated bilingual populations, as well as the development of code-switches over acquisition for the non-balanced bilingual models. Third, we explored how the patterns of the simulated groups compare with code-switching patterns extracted by two corpora that contain spontaneous utterances from Spanish–English bilingual populations.

Having established that the model reliably produces code-switched sentences, we argue that it can be employed to explain the role of syntax and semantics in specific code-switching phenomena. As a case in point, in Tsoukala et al. (in press), we employed the model to shed light on a well-known effect of verb aspect on Spanish-to-English code-switch probability. The current study's results show that the model can also account for differences in the code-switching patterns between balanced and non-balanced bilinguals.

Data Availability The script for analyzing the Miami corpus can be found at https://osf.io/vd3wa/ under miami-corpus-analysis.

Code Availability The Bilingual Dual-path model used in these experiments can be found at https://osf.io/vd3wa/ under dual-path-model.

Compliance with Ethical Standards

Conflict of Interest The authors declare that they have no conflict of interest.

References

Balam, O., Parafita Couto, M.D.C., Stadthagen-González, H. (2020). Bilingual verbs in three Spanish/English code-switching communities. https://doi.org/10.1177/1367006920911449.

Bentahila, A. (1983). Motivations for code-switching among Arabic-French bilinguals in Morocco. *Language & Communication*, *3*(3), 233–243.

Blokzijl, J., Deuchar, M., Parafita Couto, M.C. (2017). Determiner asymmetry in mixed nominal constructions: the role of grammatical factors in data from Miami and Nicaragua. *Languages*, *2*(4), 20.

Broersma, M., & De Bot, K. (2006). Triggered codeswitching: a corpus-based evaluation of the original triggering hypothesis and a new alternative. *Bilingualism: Language and Cognition*, *9*(1), 1–13.

Bullock, B.E., & Toribio, A.J. (2009). *Themes in the study of code-switching*. Cambridge: Cambridge University Press.

Chang, F. (2002). Symbolically speaking: a connectionist model of sentence production. *Cognitive Science*, *26*, 609–651.

Chang, F. (2009). Learning to order words: a connectionist model of heavy NP shift and accessibility effects in Japanese and English. *Journal of Memory and Language*, *61*(3), 374–397.

Chang, F., Dell, G.S., Bock, K. (2006). Becoming syntactic. *Psychological Review*, *113*(2), 234.

Chang, F., Baumann, M., Pappert, S., Fitz, H. (2015). Do lemmas speak German? A verb position effect in German structural priming. *Cognitive Science*, *39*(5), 1113–1130.

Deuchar, M., Davies, P., Herring, J., Couto, M.C.P., Carter, D. (2014). Building bilingual corpora. In E. Thomas, & I. Mennen (Eds.) *Advances in the study of bilingualism* (pp. 93–111). Bristol: Multilingual Matters.

Dewaele, J.-M., & Li, W. (2014). Intra-and inter-individual variation in self-reported code-switching patterns of adult multilinguals. *International Journal of Multilingualism*, *11*(2), 225–246.

Elman, J.L. (1990). Finding structure in time. *Cognitive Science*, *14*(2), 179–211.

Fishman, J.A., Cooper, R.L., Newman, R.M. (1971). *Bilingualism in the barrio* (Vol. 1). Bloomington: Indiana University.

Fitz, H., & Chang, F. (2017). Meaningful questions: the acquisition of auxiliary inversion in a connectionist model of sentence production. *Cognition*, *166*, 225–250.

Fricke, M., & Kootstra, G.J. (2016). Primed codeswitching in spontaneous bilingual dialogue. *Journal of Memory and Language*, *91*, 181–201.

Genesee, F., Nicoladis, E., Paradis, J. (1995). Language differentiation in early bilingual development. *Journal of Child Language*, *22*(3), 611–631.

Glorot, X., & Bengio, Y. (2010). Understanding the difficulty of training deep feedforward neural networks. In *Proceedings of the thirteenth international conference on artificial intelligence and statistics* (pp. 249–256).

Gollan, T.H., & Ferreira, V.S. (2009). Should I stay or should I switch? A cost–benefit analysis of voluntary language switching in young and aging bilinguals. *Journal of Experimental Psychology: Learning, Memory, and Cognition*, *35*(3), 640.

Grosjean, F. (1997). Processing mixed language: issues, findings, and models. In *Tutorials in bilingualism: psycholinguistic perspectives* (pp. 225–254).

Guzzardo Tamargo, R.E., & Dussias, P.E. (2013). Processing of Spanish-English code-switches by late bilinguals. In *Proceedings of the annual Boston University conference on language development*, (Vol. 37 pp. 134–146). NIH Public Access.

Heath, J.G. (1984). Language contact and language change. *Annual Review of Anthropology*, *13*(1), 367–384.

Janciauskas, M., & Chang, F. (2018). Input and age-dependent variation in second language learning: a connectionist account. *Cognitive Science*, *42*, 519–554.

Lantto, H. (2012). Grammatical code-switching patterns of early and late Basque-Spanish bilinguals. *Sociolinguistic Studies*, *6*(1), 21.

Lederberg, A.R., & Morales, C. (1985). Code switching by bilinguals: evidence against a third grammar. *Journal of Psycholinguistic Research*, *14*(2), 113–136.

Malhotra, S. (1980). Hindi-English, code switching and language choice in urban, uppermiddle-class Indian families. *Kansas Working Papers in Linguistics*, *5*, 39–46.

Matras, Y. (2013). Languages in contact in a world marked by change and mobility. *Revue Française de Linguistique Appliquée*, *18*(2), 7–13.

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

Myers-Scotton, C. (1993). Common and uncommon ground: social and structural factors in codeswitching. *Language in Society*, *22*, 475–503.

Petersen, J. (1988). Word-internal code-switching constraints in a bilingual child's grammar. *Linguistics*, *26*, 479–493.

Pfaff, C.W. (1979). Constraints on language mixing: intrasentential code-switching and borrowing in Spanish/English. *Language*, *55*(2), 291–318.

Poplack, S. (1978). Syntactic structure and social function of code-switching, Vol. 2. Centro de Estudios Puertorriqueños, City University of New York.

Poplack, S. (1980). Sometimes I'll start a sentence in Spanish y termino en Español: toward a typology of code-switching. *Linguistics*, *18*(7–8), 581–618.

Romaine, S. (1986). The syntax and semantics of the code-mixed compound verb in Panjabi/English bilingual discourse. In *Language and linguistics: the interdependence of theory, data and application* (pp. 35–49).

Sert, O. (2005). The functions of code-switching in ELT classrooms. *The Internet TESL Journal, 11*(8) http://iteslj.org/Articles/Sert-CodeSwitching.html.

Tsoukala, C., Frank, S.L., Broersma, M. (2017). "He's pregnant": simulating the confusing case of gender pronoun errors in L2. In *Proceedings of the 39th annual conference of the cognitive science society* (pp. 3392–3397).

Tsoukala, C., Frank, S.L., van den Bosch, A., Valdés Kroff, J., Broersma, M. (2019). Simulating Spanish-English code-switching: El modelo está generating code-switches. In *Proceedings of the workshop on cognitive modeling and computational linguistics* (pp. 20–29). Minneapolis: Association for Computational Linguistics.

Tsoukala, C., Frank, S.L., van den Bosch, A., Valdés Kroff, J., Broersma, M. (in press). Modeling the auxiliary phrase asymmetry in code-switched Spanish–English. *Bilingualism: Language and Cognition*.