Modeling change in contact settings: A case study of phonological convergence

Katia Chirkova, Tao Gong

Modeling change in contact settings: A case study of phonological convergence

Chirkova, Katia (Centre National de la Recherche Scientifique, Centre de Recherches Linguistiques sur l' Asie Orientale) & Tao Gong (Haskins Laboratories, New Haven CT, USA, Center for Linguistics and Applied Linguistics, Guangdong University of Foreign Studies, Guangzhou, China)

Abstract

Convergence is an oft-used notion in contact linguistics and historical linguistics. Yet it is problematic as an explanatory account for the changes it represents. In this study, we model one specific case of convergence (Duoxu, an endangered Tibeto-Burman language with 9 last speakers) to contribute to a more systematic understanding of the mechanisms underlying this phenomenon. The goals are (1) to address the role of some linguistic and social factors assumed to have an effect on the process of convergence, and (2) to test the following explanations of empirical observations related to phonological convergence: (a) the loss of phonological segments in a language that has undergone convergence is correlated with the relative frequency and markedness of these segments in the combined bilingual repertoire, and (b) widespread bilingualism is a prerequisite for convergence. The results of our agent-based simulation affirm the importance of frequency and markedness of phonological segments in the process of convergence. At the same time, they suggest that the explanation related to widespread bilingualism may not be valid. Our study suggests computer simulations as a promising tool for investigation of complex cases of language change in contact settings.

Convergence, agent-based modeling, bilingualism, phonology, markedness, social factors

1 Introduction: the problem

In this research paper, we use computer simulation to study phonological convergence. Convergence refers to a process of two or more languages becoming more similar in structure through contact (e.g. Hock, 1991: 492). Depending on the relationship between the languages in contact, convergence may be bidirectional (in a relationship of equal prestige) or unidirectional (in a relationship of unequal prestige). The former type refers to the mutual accommodation between languages leading to increasing equivalence of structure, as observed in linguistic areas. The latter type refers to the adjustment of a particular set of features in one language to match those of its contact language, as observed in language attrition situations. Studies in convergence distinguish between several types of convergence, such as morphological, syntactic, and phonological; the last type receives the least attention in the literature. Phonological convergence is mostly discussed as a unidirectional type of change (e.g. Clyne, 2003: 104–105, 115–116; Winford, 2003: 54–56; Matras, 2009: 224–225; Muysken, 2010: 273). Consequently, we take phonological convergence to refer to the adjustment of the inventory of sounds in the home language (that is, the one undergoing convergence) to match those of a contact language.

As convergence encompasses a broad range of changes (phonological, morphological, syntactic) associated with language contact, it is crucial to our understanding of language change in contact settings. At the same time, given that convergence can only be diagnosed after the fact, the notion of convergence merely describes the result of language change and not the process that brings about the result (see Kouwenberg, 2001). Hence, the mechanisms whereby converging varieties are created are not well understood, and the same goes for the precise factors that lead to convergence. In sum, convergence is essentially a descriptive notion, which is in need of reinterpretation and reevaluation. This paper is a preliminary step in that direction.

Our method consists of modeling one specific case of phonological convergence in the Duoxu language (ISO 639-3 code *ers*; Tibeto-Burman), whose phonological development was influenced by intensive contact with Southwestern Mandarin (hereafter, SWM, ISO 639-3 code *cnm*; Sinitic). Duoxu is part of a close-knit cluster of three languages (Duoxu, Lizu, Ersu) with a very recent split time from a shared common ancestor (Sun, 1983; Wang, 2010).[1] The close relationship between the three languages is attested by the large number of cognates with similar morphosyntactic properties and a high degree of regularity of correspondence across cognate sets (Yu, 2012; Chirkova, 2014). Starting from the 18th century, Duoxu has come in intense contact with SWM (see Sichuan Sheng Mianning Xian Difangzhi Bianzuan Weiyuanhui, 2009 for details). This contributed to divergent developments between Duoxu and its sister languages Lizu and Ersu (Chirkova, 2014). Differences in the consonant inventories are one clear example. Compared to Lizu and Ersu, the consonant system of Duoxu is considerably reduced. Notably, all consonant phonemes that are missing in Duoxu, while being present in its sister languages, involve phonological segments that are not present in SWM. Put differently, the contact influence of SWM led to the adjustment of the consonantal inventory of Duoxu to match that of SWM more closely, hence evidencing a case of phonological convergence.

The goal of our simulations is to reproduce the process of phonological convergence in the consonant system of Duoxu under the contact influence of SWM. In our simulations, we take the synchronically attested consonant system of Duoxu to represent the result of convergence between Duoxu and SWM, and the synchronically attested consonant system of SWM to represent the contact variety in the process of convergence.[2] In the absence of a complete reconstruction of the shared common ancestor of Duoxu, Lizu, and Ersu, we take the synchronically attested consonant system of Lizu, which is geographically adjacent to Duoxu, to represent the consonant system of Duoxu prior to convergence. This is admittedly a simplification. Nevertheless, given (a) a relatively recent split between Duoxu and Lizu and (b) generally regular one-to-one correspondences for consonants in these two languages (as in Table 1 below; see also Chirkova, 2014: 132–146), it is felt to be a reasonable approximation.

Computational modeling of the process of phonological convergence in Duoxu allows us to test a number of hypotheses about language-internal and language-external factors often claimed to be relevant to convergence (see in detail below). By basing our model and the simulation on real-world data (consonant inventories and consonant token counts in Duoxu, Lizu, and SWM), we can compare our simulation results with the attested data, thus assessing the relative importance of the tested factors. Naturally, simulation of real-world processes comes at the cost of simplification of the original phenomena. In our study, we use a simplified model encompassing only a selection of possible factors related to convergence. Nonetheless, it allows us to test preliminary hypotheses concerning factors related to convergence and to lay grounds for more complex and comprehensive models.

The remainder of the paper is structured as follows. Section 2 provides an overview of some fundamental assumptions related to convergence in contact linguistics and historical linguistics. It also presents the research questions investigated in our study. Section 3 details

[1] Duoxu, Lizu, and Ersu are currently classified as dialects of the same language, Ersu (e.g. Sun, 1983, 2001), and, for that reason, share the same ISO 639-3 code (*ers*). Given that Duoxu, Lizu, and Ersu are not mutually intelligible, we regard them as separate languages rather than dialects of one language (cf. Yu, 2012: 1).

[2] The phonological history of Mandarin Chinese, including its important regional variety, SWM, is well documented and understood (e.g. Norman, 1999; Yuan, 2001 and references therein). Notably, SWM is the most homogeneous of all Mandarin groups, whose consonant system, as attested synchronically, has not undergone major change since the 18th century when Duoxu came in contact with SWM.

the consonantal inventories of the three languages: Duoxu, Lizu, and SWM. Section 4 describes the major components of the model. Section 5 summarizes the simulation results. Finally, Section 6 concludes our research and points out avenues for future work.

2 Assumptions related to convergence and research questions of the study

The key assumptions associated with the notion of convergence in contact linguistics and historical linguistics include:

(1) **Source of convergence:** The linguistic behavior of bilingual speakers is the source of convergence. This is because bilinguals do not organize their communication in the form of two "languages" or "linguistic systems," but rather combine the two in one repertoire (Weinreich, 1974 [1953]; Grosjean, 1982, 1989, 2004; Matras, 2009: 224–225).

(2) **Characteristic changes associated with convergence:** Convergence consists of the leveling of differences between the languages in contact (e.g. Weinreich, 1974 [1953]: 3). In cases of phonological convergence (e.g. Dressler, 1972; Andersen, 1982; Campbell and Muntzel, 1989), phonological distinctions in the home language characteristically reduce in number, whereas distinctions shared by the home language and the contact language remain.

(3) **Factors leading to convergence:** Much like any other language contact phenomena, convergence is held to be constrained by multiple internal and external factors. Existing studies of phonological convergence mostly focus on external (social) factors relevant to the process. These include (a) social norms, awareness of identity, loyalty toward the group associated with the home language (which demand conformity to the established pronunciation norms and hence inhibit convergence), and (b) prestige of the contact language or amount of "cultural pressure" (which, on the other hand, enhance convergence) (e.g. Matras, 2009: 223–226). Internal (linguistic) factors relevant to the process of convergence receive less attention in the literature. From existing overviews of characteristic phonological changes associated with convergence (as outlined above), we infer that the loss of phonological distinctions in the home language characteristically arises through the loss of those phonological segments that do not occur in the contact language and are otherwise marked or have low frequency of occurrence (cf. Andersen, 1982: 97).[3]

(4) **Widespread bilingualism as a prerequisite for convergence:** Convergence studies suggest that widespread and stable bilingualism is an indispensable prerequisite for convergence (e.g. Weinreich, 1974 [1953]: 6; Hock, 1991: 491–493; Thomason and Kaufman, 1988: 50, 91–97; Matras, 2009). This implies that, in order for convergence changes to take effect in a population group, that group needs to have a sizeable number of bilinguals. Intuitively, this appears to be correct: if bilinguals are the agents of change, they need to be in the majority to bring about that change. However, contrary to this intuitive view, individual case studies demonstrate that bilingualism need not be extensive, and that a few bilinguals are sufficient to bring about convergence changes in a population group (e.g. Foley, 1986: 25; Rabus, 2014). Therefore, the issue of the number of bilinguals required for convergence to take effect remains controversial, and to our knowledge, it has not yet been subjected to systematic analysis.

Our study relates to all four assumptions associated with the notion of convergence, as summarized above:

---

[3] In this study we follow the classic definition of markedness as developed in the Prague School, referring to features that contribute to phonemic (binary) oppositions. We take markedness to be at the same time language-specific (that is, conditioned by language-specific systems of phonological oppositions) and correlated with a number of cross-linguistic characteristics, such as lower frequency of use, increased effort required for production, relatively greater perceptual salience, and tendency to be the target of diachronic processes or language contact processes (see Haspelmath, 2006; Bybee, 2010).

(1) We subscribe to the basic premise in convergence studies that the linguistic behavior of bilinguals is the source of convergence, as convincingly argued in the literature (see references above). Consequently, our simulations do not reflect mechanisms by which change is generated. Instead, we focus on the propagation of changes in the population undergoing convergence.

(2) We base this research on a case study (Duoxu) that exhibits characteristic changes associated with convergence (see Section 3 for details).

(3) By comparing results of our simulations of the observed set of convergence changes with real-world data, we assess the relative importance of the following language-internal and language-external factors in the process of convergence:

(a) language-internal factors: loss of phonological segments of the home language may be correlated with the relative frequency of occurrence and markedness of these segments in the bilingual repertoire
(b) language-external factors: the social factors of prestige and amount of "cultural pressure" of the contact language enhance phonological convergence

Given that our simulations only focus on those factors that enhance convergence and do not consider factors that inhibit convergence, it is expected that the endpoint of our simulations will be a complete phonological convergence between the emergent languages and SWM. The consonant system of Duoxu used in our simulations represents one stage in this process. Our focus on a selection of factors relevant to convergence (as above) allows us to clearly identify their contributions to the process.

(4) Finally, we test the intuitively plausible assumption that, for convergence changes to take effect in a population group, that group needs to have a sizeable number of bilinguals.

3 Differences between the consonant systems of Duoxu, Lizu, and SWM

The starting point for our study is the phonological system of Duoxu, as consistent with descriptions of that language based on data collected during the 1990s (Huang and Yin, 2012; Chirkova, 2015). Its consonant system comprises 33 simple consonant phonemes. The consonant inventory of Duoxu is, on the one hand, smaller than that of its sister language Lizu with 39 simple consonant phonemes (Chirkova, 2017), and, on the other hand, larger than that of SWM with 24 simple consonant phonemes (Li, 2010) (see Fig. 1 and Table A1 in the Supplementary Materials). Those consonant phonemes that occur in Lizu, but do not occur in Duoxu include /ɬ, ɹ, ɢ, q, qʰ, ɦ̃/. One more Lizu phoneme (/ɣ/) is marginal in Duoxu: /ɣ/ has low frequency of occurrence in Duoxu and only occurs before /a/ (e.g. /ɣa^{31}/ 'needle'). Table 1 details Duoxu correspondences of the Lizu phonemes /ɬ, ɹ, ɣ, ɢ, q, qʰ, ɦ̃/.[4]

Table 1. Duoxu correspondences of Lizu phonemes /ɬ, ɹ, ɣ, ɢ, q, qʰ, ɦ̃/.

No	Lizu	Duoxu	Gloss	Example	
				Lizu	Duoxu

[4] Those consonant phonemes that occur in Duoxu, but do not occur in SWM, include voiced obstruents (stops, affricates, fricatives: b, d, g, dz, dʒ, dʑ, z, ʑ, ɣ). A survey of all remaining speakers of Duoxu (Chirkova, 2014) reveals that some of them have already lost a distinction between voiceless and voiced initials, having replaced voiced initials by their voiceless counterparts. This development, correlated with the present, critically endangered state of the language, represents a complete convergence between the consonant systems of Duoxu and SWM. Note that convergence in one subsystem of a language (e.g. phonology) cannot be equated with a complete language shift; speakers can perfectly well maintain their language (the lexicon and grammar) while the phonological system of that language may undergo change. This is exactly what we observe in Duoxu.

1	ɬ	l	‘poplar’	/ɬæ^{33}ɬæ^{33}se^{53}/	/lo^{31}lo^{53} se^{53}/
2	ɹ	w	‘to swell, to bulge’	/[de^{55}]-ɹɐ^{33}[-ɹɐ^{31}]/	/wa^{22}/
3	ɣ		‘needle’	/ɣɐ^{53}/	/ɣa^{31}/
4	ɢ	g	‘plate’	/ɴɢɐ^{55}mɐ^{51}/	/ga^{33}ma^{33}/
5	q	k	‘steelyard’	/qɐ^{51}/	/ke^{22}/
6	qʰ	kʰ	‘walnut’	/qʰɐ^{55}ɹə^{31}/	/kʰa^{33}wu^{33}/
7	h̃	m	‘bamboo’	/h̃e^{55}/	/mi^{33}/
		n	‘year’	/h̃e^{55}/	/nje^{33}/
		ŋ	‘to smell good, fragrant’	/[de^{33}-]h̃jo^{53}/	/ŋo^{33}/

The set of changes between Lizu and Duoxu is consistent with the characteristic phonological changes associated with phonological convergence, as outlined above. First, all Lizu phonemes in Table 1 do not occur in SWM, the contact language of Duoxu (see Fig. 1, where Lizu is displayed as pre-convergence Duoxu – Duoxu_PCVG – as explained in Section 1). Second, with the only exception of the correspondence /ɹ/ to /w/, all correspondences in Table 1 represent the substitution in Duoxu of the marked member of an opposition in Lizu (e.g. /ɬ/ vs. /l/, uvular stops vs. velar stops, /h̃/ vs. modal voiced nasals) by its unmarked member. In sum, the set of changes in Table 1 clearly illustrates the reduction of the overall number of consonantal segments in Duoxu as a result of the loss of marked segments to yield a consonant system that is more in line with the SWM phonology.

The changes in Table 1 are likely to have occurred rapidly sometime after Duoxu came in close contact with SWM. Most of them are already reflected in the earlier attestations of Duoxu in Chinese and Tibetan transcription, dating from the mid-18th century (Nishida, 1973; Chirkova, 2014).[5] While no attestations of Lizu from the same time period are available for comparison, we note that earlier descriptions of Lizu by Chinese linguists dating from the 1980s (Huang and Renzeng, 1991) document a rich consonant system including all of the phonemes listed in Table 1. The same consonant phonemes are also attested across various subvarieties of Lizu (cf. Yu, 2012), indicating their conservative nature. It is therefore reasonable to assume that, unlike Duoxu, the consonant system of Lizu remained relatively stable through the reference period (from the beginning of the 18th century until the present time).

In our study, both the set of convergence changes and the complex consonant inventories of Duoxu and Lizu are simplified to make them more amenable to computer modeling and analysis. First, among the changes listed in Table 1, we only focus on one-to-one correspondences between Lizu and Duoxu (Nos. 1 to 6). The one-to-many correspondence (No. 7), which reflects the opposition between voiced and voiceless nasals in Lizu, has not been taken into account.[6] In other words, the original consonant inventory of Lizu, consisting of 39 consonant phonemes, is reduced to 38 phonemes.

Second, to facilitate the comparison between the consonant inventories of the three languages, we include one allophone of a Lizu phoneme: [f]. [f] is an allophone of /x/ in Lizu, but it is an independent phoneme in Duoxu and SWM. The list of consonants compared between the three languages hence comprises a total of 39 consonants (38 phonemes and one allophone).

[5] It remains a possibility that those transcriptions may be biased towards the production of those bilingual speakers who show stronger effects of convergence than an average L1 Duoxu speaker.

[6] The frequency count of Lizu /h̃/ (a total of 30 tokens) is incorporated into the counts of Lizu nasals: 7 tokens are added to the total number of tokens for /m/; 8 tokens to the total number of tokens for /n/; 7 tokens to the total number of tokens for /ɲ/; and 8 tokens to the total number of tokens for /ŋ/.

Third, in our simulations we only focus on simple consonant phonemes of Lizu and Duoxu and do not take into account complex consonant phonemes (consonant clusters) in the two languages. Overall, the consonant cluster system in Lizu is more complex, whereas most corresponding clusters have been simplified in Duoxu. Examples include: ‘plate’: Lizu /ɴɢɐ^{55}mɐ^{51}/, Duoxu /ɡa^{33}ma^{33}/. Much like the set of changes in Table 1, simplification of the Duoxu consonant cluster system (including homorganic nasal clusters /mb, nd, ŋɡ/) is an example of the reduction of the overall number of phonological segments in Duoxu, while preserving segments common to both Duoxu and SWM. However, unlike the set of changes in Table 1, the process of simplification in the consonant cluster system of Duoxu is not yet complete. In addition, sound correspondences between consonant clusters in Duoxu and Lizu are more complex than those between simple consonant phonemes.[7] For the purpose of our simulations, we simplify the consonant systems of Lizu and Duoxu to only contain simple consonants; more specifically, we calculate the roots of consonant clusters together with their corresponding simple consonant phonemes. To give one example, the total number of tokens for Lizu /ɡ/ combines the token numbers for both /ɡ/ and /ŋɡ/.

Finally, our simulations are based on the natural occurrence frequencies of consonants in Duoxu, SWM, and Lizu in the same corpus of 1260 basic words (based on both published and firsthand fieldwork data: Li, 2010, p.c.; Chirkova, 2014). The precise token count is provided in Table 1 in the Supplementary Materials.

Figure 1 provides a comparison of consonant frequencies in (a) Lizu, taken to approximate a pre-convergence stage of Duoxu (Duoxu_PCVG) before the set of substitution changes in Table 1 occurred, (b) Duoxu, representing the post-convergence state of the language, and (c) SWM, the contact language.

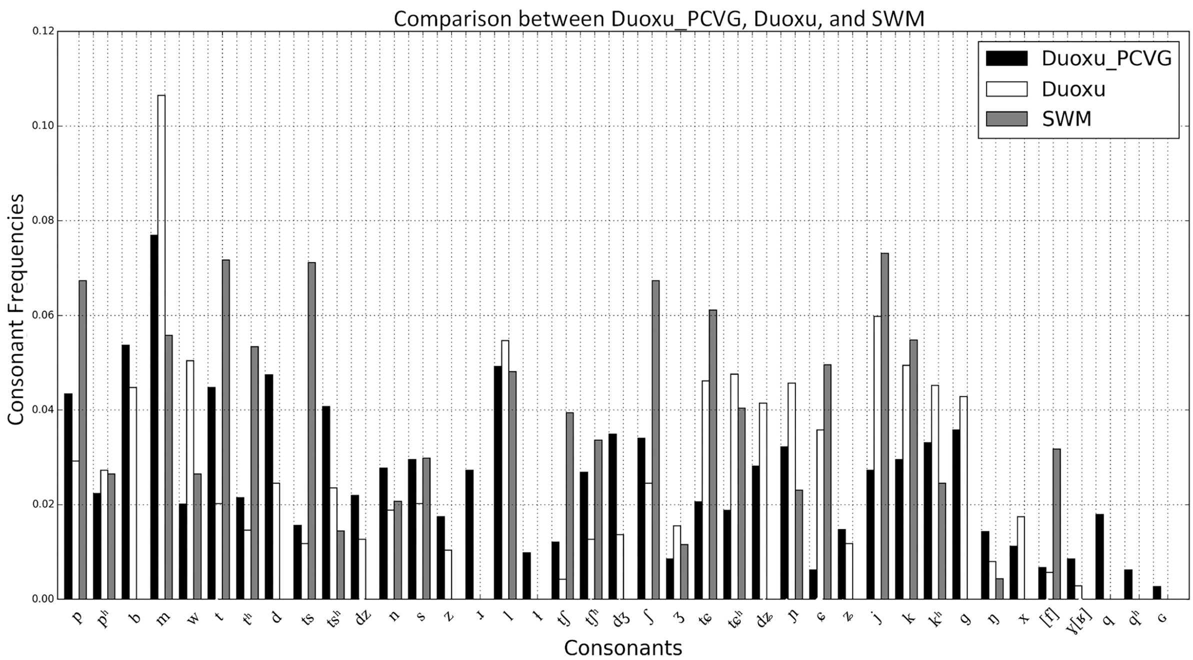


Figure 1. Consonant frequencies in Duoxu_PCVG, Duoxu, and SWM. The plot is drawn on the basis of the normalized consonant frequencies in Table 1 in the Supplementary Materials.

[7] Of all consonant clusters, the Lizu cluster /mb/ has the most complex correspondence patterns in Duoxu. Lizu /mb/ may correspond to Duoxu: (1) /mb/, as in ‘mountain’: Lizu /mbje^{23}/, Duoxu /mba^{33}/; (2) /m/, as in ‘horse’: Lizu /mbɹə^{53}/, Duoxu /mo^{31}/; and (3) /b/, as in ‘deaf person’: Lizu /nɐ^{33}mbo^{53}/, Duoxu /njo^{22}bo^{33}/.

A comparison of consonant frequencies in Lizu and Duoxu in Fig. 1 reveals two major types of changes between these two languages:

(1) the loss in Duoxu of the consonants /ɬ, ɹ, ɣ, ɢ, q, qʰ/, and their substitution by the consonants /l, w, g, k, kʰ/; and

(2) changes in the frequency of occurrence of those consonants that are shared by Duoxu_PCVG and Duoxu, most importantly an increase in the frequency of the Duoxu consonants /l, w, g, k, kʰ/. These changes are likely to be a direct consequence of the convergence changes affecting the consonants /ɬ, ɹ, ɣ, ɢ, q, qʰ/.

These two types of changes will be used in assessing similarity of the emerging language varieties in our simulations with Duoxu.

4 Model architecture

The Python code for the simulations and figures, the R code for the statistical analyses of the simulation results, and the csv files of the consonant distributions in Lizu (Duoxu_PCVG), Duoxu, and SWM can be found at https://github.com/gtojty/ConsEvo.

4.1 Agents and language groups

Language contact in our study is modeled as a process of consonant exchange in a population of artificial agents. Our choice of an agent-based model is motivated by the complexity of the phenomenon of convergence and the resulting (1) difficulty of defining mutual influence among different sets of consonants through mathematical equations, (2) necessity to focus on available knowledge of individual behaviors rather than on group-level activities, and (3) need of stochasticity (an important property of agent-based models) for the modeling of these behaviors.

As in many multi-agent models of language evolution, the agents in our model are small computer programs that possess certain characteristics and operate in an autonomous fashion according to some predefined or gradually developed regulations, without much human interference (Steels, 1997; Ferber, 1998; Gong et al., 2014). As language users, agents can preserve phonemic contrasts among the consonants in question, and produce, perceive, and store these consonants in their repertoires (memory). Sharing relevant consonants among agents makes contact possible; when 'hearing' any of these consonants, agents can remember and produce them in future communications. Throughout the simulation, agents are not allowed to introduce new consonants; in other words, our model simulates a closed system that focuses on consonant exchange between Duoxu_PCVG and SWM.

In the consonant repertoire of an agent, each consonant is stored as a distinct prototype, together with a value recording its frequency of occurrence in communications. The frequency of occurrence of a consonant reflects the proportion of the use of that consonant among all other consonants perceived by an agent in previous communications. These frequency values are updated during communications when the agent acts as a listener.

At the beginning of a simulation, agents are assigned to one of the three language groups:

(1) Monolingual SWM speakers, who are initialized with the frequency of occurrence of SWM consonants.

(2) Monolingual Duoxu_PCVG speakers, who are initialized with the frequency of occurrence of Duoxu_PCVG consonants.

(3) Duoxu_PCVG-SWM bilingual speakers, who are initialized with the accumulated frequency of occurrence of the consonants in SWM and Duoxu_PCVG (the accumulated frequencies are obtained by adding the normalized frequencies of Duoxu_PCVG and SWM and then normalizing the summed frequencies). This setting is taken to reflect the findings in

the literature on bilingual language processing that (a) phonological systems in bilinguals are likely to be merged and mutually influence each other and that, more generally, (b) bilinguals activate both languages in comprehension and production (even when engaged in monolingual tasks) (e.g. Guion, 2003; Desmet and Duyck, 2007; Antoniou et al., 2011; Kroll et al., 2015).[8]

Group 1 forms the SWM-speaking population, while groups 2 and 3 form the Duoxu_PCVG speaking population. Contact between the SWM and Duoxu_PCVG populations is brought about by bilingual Duoxu_PCVG-SWM speakers in group 3.

4.2 Communications

Communications, also called language games in many computational models of language evolution (Steels, 1997; Gong et al., 2014), are pairwise interactions between two agents (a speaker and a hearer) who are chosen from the entire population randomly and in accordance with the ratios of monolinguals and bilinguals in the population (see Section 4.4). This is admittedly a simplification, compared to real-life communications; however, it does not alter the principles discussed below.

Given the main focus of our study on phonological convergence in the population of Duoxu_PCVG speakers, our model addresses the following four types of communications involving (bilingual and/or monolingual) individuals from that population:[9]

(1) A Duoxu_PCVG-SWM bilingual talks to another Duoxu_PCVG-SWM bilingual, using the combined repertoire of Duoxu_PCVG and SWM consonant phonemes.

(2) A Duoxu_PCVG monolingual talks to another Duoxu_PCVG monolingual, using consonants from the consonant inventory of Duoxu_PCVG.

(3) A Duoxu_PCVG monolingual talks to a Duoxu_PCVG-SWM bilingual or a Duoxu_PCVG-SWM bilingual talks to a Duoxu_PCVG monolingual. In this type of communications, a Duoxu_PCVG monolingual uses consonants from the consonant inventory of Duoxu_PCVG, whereas a Duoxu_PCVG-SWM bilingual uses the combined repertoire of Duoxu_PCVG and SWM consonant phonemes.

(4) A SWM monolingual talks to a Duoxu_PCVG-SWM bilingual, using consonants from the consonant inventory of SWM.

We exclude those types of communications in which SWM monolinguals act as listeners (that is, SWM monolinguals talking to other SWM monolinguals, or Duoxu_PCVG-SWM bilinguals talking to SWM monolinguals). Such communications have no direct influence on the language of the Duoxu_PCVG population and are therefore irrelevant for our purposes.

Among the four types of communications, the first three are bidirectional, that is to say, both agents participating in those types of communications can act as either speakers or listeners. The last type (in which SWM monolinguals talk to Duoxu_PCVG-SWM bilinguals) is unidirectional. In this type of communications, SWM monolinguals can only act as speakers, whereas Duoxu_PCVG-SWM bilinguals can only act as listeners. The ratio of the four types

---

[8] For the purpose of the present investigation, we assume that bilinguals do not distinguish between the two languages regarding the frequency with which phonemes occur. Another, more nuanced possibility involves considering two internal representations for bilinguals that may or may not influence each other. We hope to address this issue in future research.

[9] These types of communications broadly correspond to Grosjean's (2004: 40) "language modes" of bilinguals: "A mode is a state of activation of the bilingual's languages and language processing mechanisms. […] At one end of the continuum, bilinguals are in a totally monolingual language mode in that they are interacting only with (or listening only to) monolinguals of one or the other of the languages they know. […] At the other end of the continuum, bilinguals find themselves in a bilingual language mode in that they are communicating with (or listening to) bilinguals who share their two (or more) languages and where language mixing may take place (i.e., code-switching and borrowing)."

of communications in the Duoxu_PCVG population is determined by the proportion of bilingual and monolingual agents in that population (see Section 4.4).

In our model, each communication proceeds in two steps: (1) production and (2) comprehension.

In the production process, the speaker produces a set of repeatable consonants, which are chosen from the consonant repertoire that is proper to the group to which that speaker belongs (see above). The probabilities with which those consonants are chosen for production are determined by the frequency of occurrence of different consonants in the speaker's repertoire.

In the comprehension process, the hearer receives the set of consonants produced by the speaker, calculates the frequency of occurrence of those consonants within that set, and adds the calculated frequency of occurrence to the corresponding consonants in the existing repertoire for that hearer. After that, the hearer normalizes the frequencies of occurrence of all consonants in its existing repertoire. The normalization operation enables (a) future production of consonants when that agent acts as a speaker and (b) calculation of consonant frequencies in the emergent variety.[10]

Our model makes no distinction between successful or failed communications, neither does it include transmission errors. Instead, communications in our model are a process of consonant sampling and exchange. The choice of the language for communication in a particular setting is determined by the language groups from which the speaker and the hearer are selected. The outcome of sampling depends on the choice of the set of consonants produced by the speaker and perceived by the hearer.[11]

4.3 Simulation scenarios

We consider the following two simulation scenarios corresponding to our research questions outlined in Section 3.

Scenario 1 (linguistic factors hypothesis): In this scenario, we test the intuitively plausible suggestion in convergence studies that the loss of consonant phonemes of the home language is correlated with their (a) low frequency of occurrence and (b) markedness in the bilingual repertoire. In relation to our case study, the assumption under this hypothesis is that the consonants /ɬ, ɹ, ɣ, ɢ, q, qʰ/ are substituted by the consonants /l, w, w, ɡ, k, kʰ/ because the former have low frequency of occurrence and are marked. These two factors (low frequency of occurrence and markedness) are implemented in our model as a markedness threshold, defined in reference to the frequency of occurrence of /ɬ, ɹ, ɣ, ɢ, q, qʰ/ in the combined Duoxu_PCVG-SWM bilingual repertoire. The markedness threshold serves to formally distinguish all phonemes in that set from all other phonemes in the combined Duoxu_PCVG-SWM repertoire by combining the properties of low frequency of occurrence with some other properties (such as increased effort required for production and relatively greater perceptual salience) that are associated with the notion of markedness cross-linguistically. This

[10] This implies the assumption that the weight of the recently heard evidence (that is, the phoneme frequencies in the utterances just heard) is equivalent to the agent's language representation, which may be a simplification. Further work may need to use a weighted update of the frequencies of occurrence of different consonants or a Bayesian model that takes into account not only the frequencies of occurrence of different consonants, but also their number.

[11] Apart from this way of manipulation, another option is to confine the probabilities for an individual to interact with different language groups, as implemented in some language evolution models (e.g. Smith and Hurford, 2003). However, this option appears to be more appropriate in situations that involve homogeneous agents (Gong, 2010). It is, therefore, less suited to the present modeling settings, which involve heterogeneous agents (monolingual and bilingual speakers, which furthermore use different languages to communicate with different agents). For this reason, we opt for directly controlling the ratios of different types of communications at the population level.

implementation reflects the observation that, while all phonemes in the set /ɬ, ɹ, ɣ, ɢ, q, qʰ/ generally have low frequency of occurrence, it is only /ɬ, ɣ, ɢ, qʰ/ that have the lowest frequency of occurrence among all phonemes in the combined Duoxu_PCVG-SWM repertoire. The frequency of occurrence of the two remaining phonemes (/ɹ, q/) is higher than that of some other phonemes that do not undergo change in Duoxu (such as /ʑ/; see Table A1). In other words, in addition to low frequency of occurrence, Duoxu phonemes in the set /ɬ, ɹ, ɣ, ɢ, q, qʰ/ are likely to be characterized by additional properties that trigger their substitution in the real-world Duoxu case, as opposed to other low-frequency phonemes. This is formally specified by the markedness threshold parameter.

Once the combined frequency values for a consonant in the set /ɬ, ɹ, ɣ, ɢ, q, qʰ/ fall below the predefined markedness threshold, that consonant is subject to substitution by the corresponding consonant in the set /l, w, w, g, k, kʰ/. (See Supplementary Materials for an analysis of the robustness of this parameter.) Consonant substitution occurs at the stage of production: if a consonant in the set /ɬ, ɹ, ɣ, ɢ, q, qʰ/ is chosen for production by a speaker, and the frequency of occurrence of that consonant in that speaker's repertoire is below the predefined threshold, that speaker will produce instead the corresponding consonant in the set /l, w, w, g, k, kʰ/. Following the substitution operation, the speaker implements a gradual adjustment of the frequency of occurrence of the produced consonant. This ensures that the substitution of the consonants set /ɬ, ɹ, ɣ, ɢ, q, qʰ/ is gradual as well, so as to reflect the view that phonological change is phonetically abrupt but lexically gradual (e.g. Wang, 1969). The update function consists of adding a predefined frequency adjustment value to the frequency of occurrence of the produced consonant and deducting the same predefined frequency adjustment value from the frequency of occurrence of the original consonant. The update function is used in all types of communications and in all scenarios in which consonant substitution can occur (more details on the update function are provided in Section 4.4). If, after this adjustment, the frequency of occurrence of the original consonant is below 0.0, that consonant is discarded. At that point, the speaker is said to have completed the substitution of that consonant.

Scenario 2 (social factors hypothesis): In this scenario, we test the assumption in convergence studies that the social factors of prestige of the contact language and/or amount of "cultural pressure" may enhance convergence. The assumption under this hypothesis is that the consonants /l, w, w, g, k, kʰ/ replace the consonants /ɬ, ɹ, ɣ, ɢ, q, qʰ/ because the former are considered to be more prestigious variants of the latter. This would reflect an often unconscious and likely unconditional tendency of speakers (either bilingual or monolingual) to adjust both linguistic and paralinguistic aspects of their speech in order to gain identification with the interlocutor's social group (e.g. Shepard et al., 2001). It is implemented as a direct, unconditional substitution of a consonant in the set /ɬ, ɹ, ɣ, ɢ, q, qʰ/ by its respective unmarked counterpart. Once a consonant in the set /ɬ, ɹ, ɣ, ɢ, q, qʰ/ is chosen for production, the speaker will directly produce the corresponding consonant in the set /l, w, w, g, k, kʰ/, and adjust the frequency of occurrence of the relevant consonants as outlined above in relation to Scenario 1.

Depending on the type of communications in which consonant substitution can occur, the two scenarios are divided into two sub-scenarios each:

Scenario 1a and scenario 2a: Consonant substitution only occurs in bilingual-bilingual communications.

Scenario 1b and scenario 2b: Consonant substitution can occur in all types of communications, involving both monolingual and bilingual Duoxu_PCVG speakers.

These four scenarios, all of which involve consonant substitution, are compared to a reference scenario (**Scenario 0**) that does not involve any consonant substitution. Table 2 summarizes the five scenarios, the factors they involve, and the types of communications in which the substitution changes can occur.

Table 2. Simulation scenarios, related factors, types of communications.

No	Name	Factors	Types of communications
0	Reference scenario	none	
1a	Linguistic factors scenario	low frequency of occurrence, markedness	bilingual-bilingual communications
1b			all types
2a	Social factors scenario	prestige, "cultural pressure"	bilingual-bilingual communications
2b			all types

4.4 Model parameters

Our model is specified by six parameters, as outlined in Table 3.

Table 3. Parameters of the model and their values in the simulations.

Label	Meaning	Value
N_P	Number of agents per population	100
R_{BI}	Ratio of bilinguals in the Duoxu_PCVG population	0.0 to 0.9
N_C	Number of communications	10,000
N_{CONS}	Number of consonants exchanged in a communication round	300
F_{MRK}	Markedness threshold for consonant substitution in scenario 1	0.01
F_{ADJ}	Frequency adjustment for consonant substitution	0.002

(1) N_P: We set the size of the two populations (Duoxu_PCVG and SWM) to 100 agents each because we do not know the exact population numbers for Duoxu and SWM groups at the time when the set of convergence changes in Duoxu occurred. This is admittedly an idealized and simplified setting. Nonetheless, it is useful for a quantitative evaluation of our hypotheses. In addition, assuming equal size of the SWM and Duoxu_PCVG populations does not affect simulation results, because the ratios of different types of communications are determined mainly by the ratio of monolingual and bilingual speakers in the Duoxu_PCVG population (our second parameter, R_{BI}, see (2)), rather than the size of the SWM population.

(2) R_{BI}: This parameter is used to control the proportion of Duoxu_PCVG-SWM bilingual speakers in the Duoxu_PCVG population. R_{BI} ranges from 0.0 to 0.9, with a step of 0.1. (The proportion of Duoxu_PCVG monolinguals accordingly equals $1–R_{BI}$ and ranges from 1.0 to 0.1.) Simulations with R_{BI}=0.0 are meant to trace the evolution of the original, pre-convergence state of the language without contact.

(3) N_C: The total number of communications involving Duoxu_PCVG agents is set to 10,000 in all simulations or 100 communications per agent. This enables a reasonable number of communications to implement consonant substitution changes in the emergent varieties and to observe the effect of the length of contact on these varieties.

(4) N_{CONS}: The number of consonants exchanged in each communication is set to 300. This value is chosen with two considerations in mind. On the one hand, the number of consonant exchanges in one communication needs to allow even the least frequent consonant to be produced at a chance level, which can be described as in Equation (1):

$$p(k) = 1 - \left(1 - p(k)\right)^N > 0.5 \qquad (1)$$

where $p(k)$ is the frequency of occurrence of the least frequent consonant. In our case study, $p(k)=p(\text{G})\sim 0.02$, so we can derive that $N > 347$. On the other hand, we need to consider the bottleneck effect in cultural transmission: in the learning of a cultural system such as language, learners may not necessarily receive all data that constitute that system (e.g. Kirby, 1999). Consequently, we would expect that not all consonants in the consonant inventory of a speaker are produced in every communication. To balance the two considerations, we reduced the mathematically calculated value of 347 to a round 300.

(5) F_{MKD}: The markedness threshold reflecting the effect of low frequency of occurrence and markedness in scenarios 1a and 1b is set to 0.01. This value is defined in reference to the frequency of occurrence of the consonants /ɬ, ɹ, ɣ, ɢ, qʰ, q/ in the combined Duoxu_PCVG-SWM bilingual repertoire (see Section 4.3 and Table A1). (The effect of varying the value of F_{MKD} is discussed in the Supplementary Materials.)

(6) F_{ADJ}: The role of this parameter is to ensure that the substitution of consonants in all scenarios where it occurs, is gradual, as motivated in Section 4.3 above. In our simulations, the gradual adjustment value is set to 0.002. This value is chosen in correlation with F_{MKD} (set to 0.01). By setting the gradual adjustment value to 0.002, we ensure that even consonants with a frequency of occurrence of 0.01 (the markedness threshold) can be chosen for production at least five times before they are substituted.[12] Note that varying the F_{ADJ} value has little effect on the simulation results (see Fig. A3 in the Supplementary Materials).

In each simulation scenario, we conduct ten sets of simulations based on the ten values of R_{BI}. In each set, we collect the results of 20 runs for analysis. In each run, we set 51 sampling points distributed evenly between 0 and 10,000 communications (100 per agent) with a step of 200 (2 per agent). At each sampling point, we record the frequency of occurrence of all consonants in the consonant repertoires of all 100 agents in the Duoxu_PCVG population (including both bilinguals and monolinguals).

4.5 Analysis of simulation results

Our analysis of the simulation results comprises two parts. First, we find among all emergent varieties those that most resemble present-time Duoxu in terms of their consonant frequencies. To this end, we use a quantitative measure called Summed Square Difference (SSD), derived from the statistical summed square error, as defined in Equation (2):

$$SSD = \sum_{i=1}^{n}(x_i - y_i)^2 \qquad (2)$$

where *n* is the number of shared consonants between the two varieties X (emergent variety) and Y (Duoxu), and x_i and y_i are the normalized frequencies of occurrence of the consonant *i* in varieties X and Y, respectively. If a consonant is missing in one language, its normalized frequency of occurrence is 0.0. By taking square of the difference, we ensure that both positive and negative differences are treated equally and that we can obtain an accumulated difference among all consonants of the two varieties. A comparison of emergent varieties with the lowest SSD values across different simulation scenarios enables us, on the one hand, to assess the respective roles of the linguistic and social factors yielding an emergent variety that resembles Duoxu, and on the other hand, to examine the role of bilinguals in the process.

Second, we directly compare the degree of implementation of the diagnostic set of changes outlined in Section 3 in the emergent varieties with the lowest SSD values to those in both Duoxu_PCVG and present-day Duoxu. This enables us to assess the range of changes undergone by each individual emergent variety in contact with SWM and the conformity of these changes to those observed between Duoxu_PCVG and present-day Duoxu.

[12] Note that the value of F_{ADJ} is also set in correlation with the total number of communications in our simulations. Overall, the lower the value of F_{ADJ}, the more communications it takes to implement all substitution changes in Table 1.

5 Simulation results

To examine the effects of the tested factors on the emergence of Duoxu-like language varieties, we employ a generalized linear regression model using the tested factors to predict the lowest SSD values. (The regression model is used to avoid an unnecessarily large number of simulations. In relation to our model, which targets only some possible factors, running many simulations entails the risk of exaggerating the effect of uncontrolled factors and leading to divergence.) The model is run using R 3.2.4 (R Core Team, 2016). The lowest SSD value is the dependent variable.[13] There are four independent variables (predicting factors), including:

(a) Ratio of bilinguals. This is a continuous variable with ten values (based on R_{BI}).

(b) Low frequency and markedness. This is a categorical variable with two levels:
- 0: markedness threshold not in effect (scenarios 0, 2a and 2b)
- 1: markedness threshold in effect (scenarios 1a and 1b)

(c) Social factors. This is a categorical variable with two levels:
- 0: social factors not in effect (scenarios 0, 1a and 1b)
- 1: social factors in effect (scenarios 2a and 2b)

(d) Types of communications in which consonant substitution can occur. This is a categorical variable with two levels:
- 0: consonant substitution occurs in all types of communications (scenarios 1b and 2b)
- 1: consonant substitution is restricted to bilingual-bilingual communications (scenarios 1a and 2a)

In each scenario, there are 20 lowest SSD values obtained in 20 simulations for analyses. The critical *p* value for identifying significant effects is set to 0.05.

The linear regression model also includes an interaction between language-internal and language-external factors (low frequency of occurrence & markedness and social factors, respectively) and types of communications.

Table 4. Results of the linear regression model. Numbers in brackets indicate the levels of the categorical variable. "*β*" represents the regression coefficient. Significant factors ($p < 0.05$) are marked in bold.

Independent factors	***β***	**Std.error**	***t***	***p***
Ratio of bilinguals	**.00063**	**.00010**	**6.367**	**< .00001**
Low frequency and markedness (1)	-.00008	.00006	-1.368	.172
Social factors (1)	**-.00144**	**.00006**	**-22.910**	**< .00001**
Types of communications (1)	**.00078**	**.00006**	**12.416**	**< .00001**
Types of communications (1) x Low frequency and markedness (1)	**-.00068**	**.00009**	**-7.641**	**< .00001**

[13] Our focus on the lowest SSD value rather than on all possible SSD values throughout all simulation runs under the same setting is motivated by the following considerations. First, the lowest SSD value enables us (a) to locate among all emergent varieties one that mostly resembles Duoxu in terms of its consonant distribution, and (b) to investigate the conditions under which this variety emerges. Given that the minimum SSD value and the precise conditions under which it is obtained may vary slightly due to random noise in different runs under the same setting, we consider 20 lowest SSD values obtained in 20 simulations for analyses (see below). Second, while an analysis of all SSD values throughout all runs under the same setting may be indicative of the general language contact dynamics in the model, we find that this dynamics is not considerably different under different R_{BI} values. Instead, the lowest SSD values are more dependent on R_{BI} and the exact simulation scenario. This is illustrated in Fig. 2 and in Fig. A1 in the Supplementary Materials. Third, the state in which varieties that most resemble Duoxu emerge is generally not stable (see below). For this reason, focusing on all SSDs cannot give us much information about those particular states.

Table 4 shows the results of the statistical model. Of all independent factors, only the factor of low frequency and markedness does not have a significant main effect on the lowest SSD values, as compared to the reference scenario (scenario 0). However, this factor interacts significantly with the factor types of communications. The precise effects of significant factors differ, as reflected by their β values. Positive β values indicate that an increase in relevant independent variables (that is, ratio of bilinguals, types of communications) leads to an increase in SSD values. Negative β values, on the other hand, indicate that, when social factors or an interaction of types of communications and low frequency and markedness are in effect, SSD values will decrease.

The effect of the ratio of bilinguals on the lowest SSD values can be observed by comparing mean SSD values obtained at different sampling points throughout all communications in different simulation scenarios and under different ratios of bilinguals. This is illustrated in Fig. 2 in relation to scenario 2b. (The evolution of mean SSD values in the remaining four scenarios 0, 1a, 2a, 1b is presented in Fig. A1 in the Supplementary Materials.)

In simulations involving bilinguals (R_{BI}=0.1 to R_{BI}=0.9), the general trend is that the higher the ratio of bilinguals, the quicker the emerging varieties will attain complete convergence with SWM, largely bypassing an intermediate state that would resemble Duoxu. By contrast, the smaller the ratio of bilinguals, the longer the emerging languages can maintain low SSD values and the greater the chance to attain an intermediate state that resembles Duoxu. In simulations without bilinguals (R_{BI}=0.0), SSD does not change much throughout simulations, regardless of the number of communications. This means that when the population of agents contains no bilinguals, the original consonant system remains largely unchanged throughout its evolution.

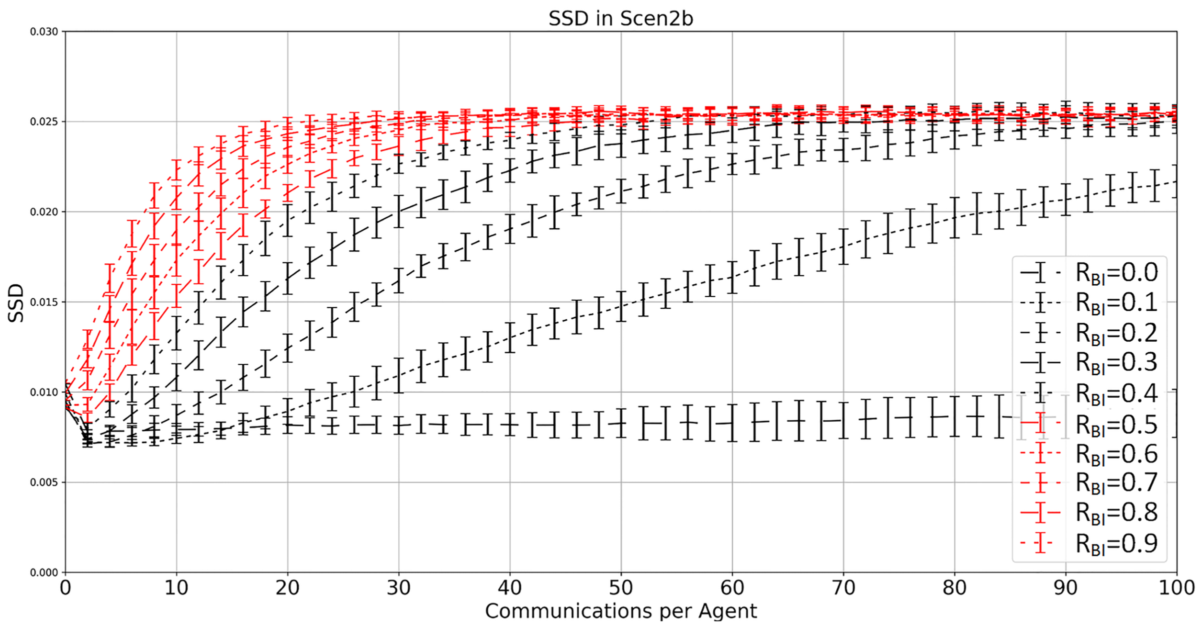


Figure 2. Mean SSD values obtained throughout 100 communications per agent in scenario 2b. Each curve corresponds to the result under a particular R_{BI}. Results are averaged over 20 runs in each condition. Error bars denote standard errors.

The endpoint of all simulations with bilinguals (10,000 communications or 100 communications per agent, R_{BI}=0.1 or higher) is an emergent variety that, in terms of the inventory and the frequency of occurrence of its consonants, is by and large identical to SWM. In other words, the endpoint of all simulations is complete convergence between Duoxu_PCVG and SWM. This is illustrated in Fig. 3 in relation to scenario 2b under R_{BI}=0.5.

Such a development is not contradictory to the historical process, as we observe complete phonological convergence between Duoxu and SWM in the speech of some Duoxu speakers, due to the present critical degree of endangerment of this language (see footnote 4).

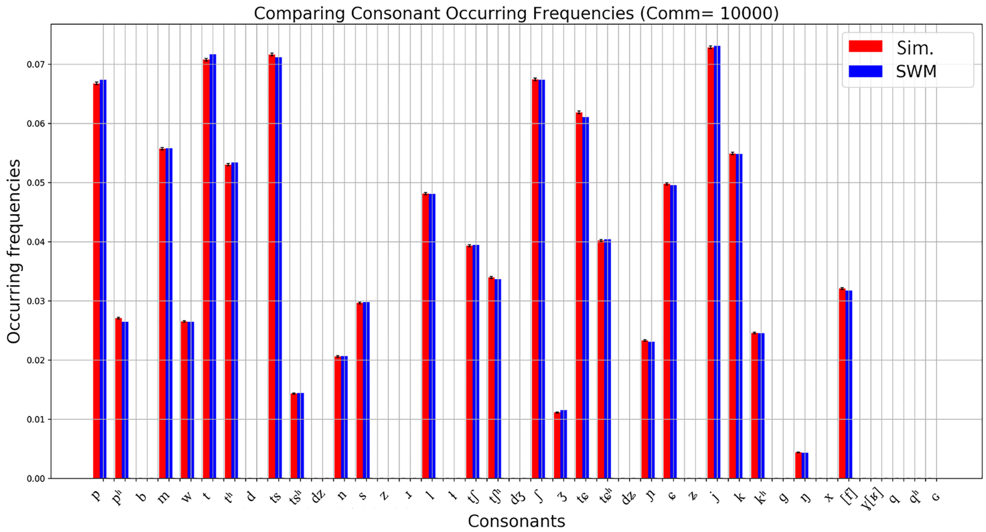


Figure 3. Comparison of the emergent variety in scenario 2b under R_{BI}=0.5 at the endpoint of the simulations (10,000 communications or 100 communications per agent) and SWM.

A comparison of the lowest SSD values under R_{BI} values from 0.0 to 0.9 and across the five scenarios used in our simulations points to the following two scenarios that are most successful in lowering SSD values: (a) scenario 1b, in which the markedness threshold is in effect and consonant substitution occurs in all types of communications, and (b) scenario 2b, in which social factors are in effect and consonant substitution also occurs in all types of communications.

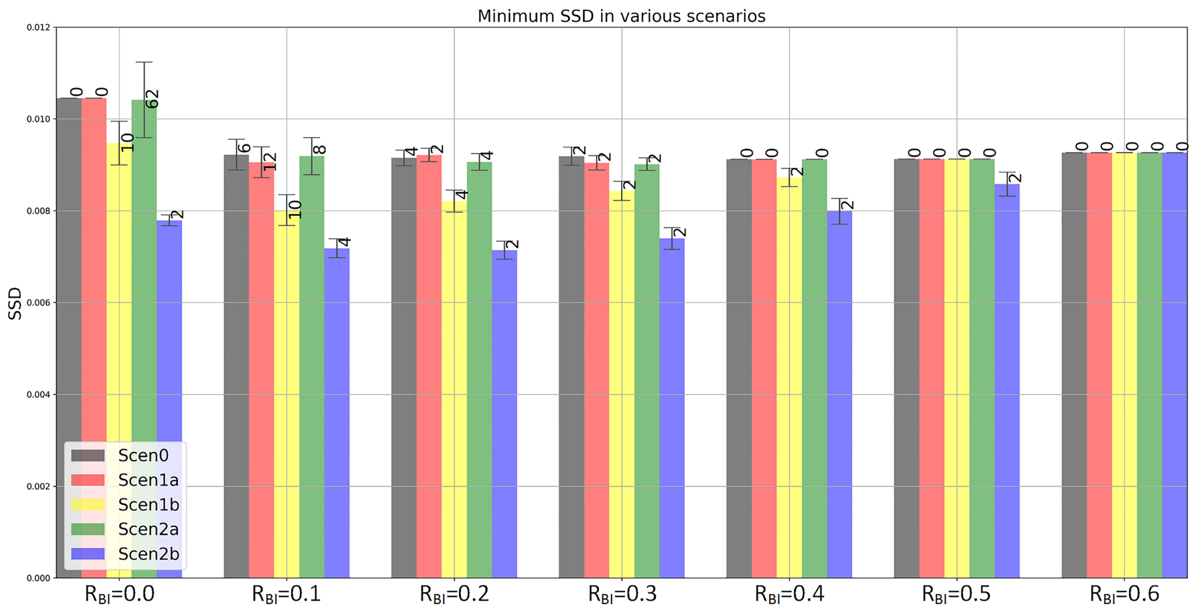
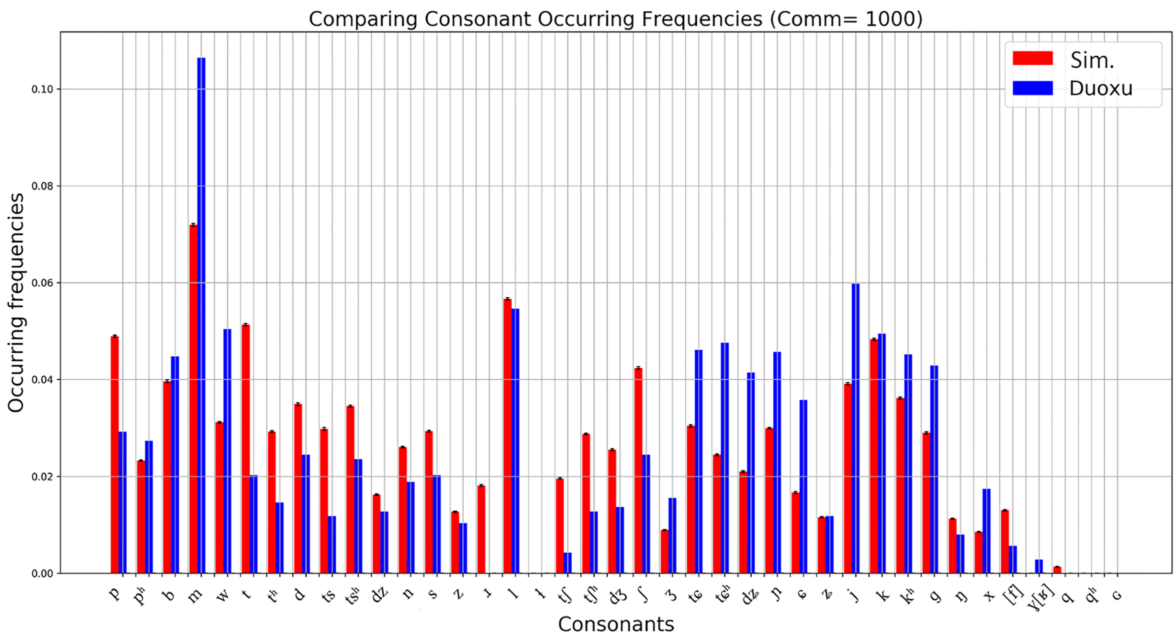


Figure 4. Lowest SSD values in different scenarios under R_{BI} values from 0.0 to 0.6. Results for R_{BI} values from 0.7 to 0.9 are omitted because they are identical to those under R_{BI}=0.6,

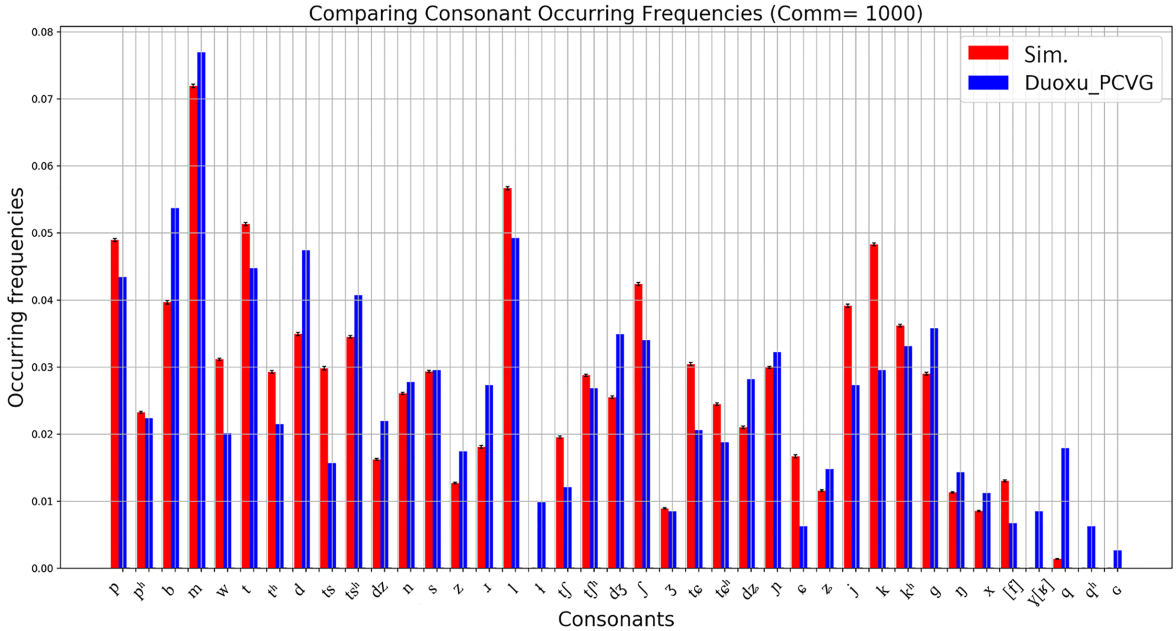
where the lowest SSD value is attained at 0 communications. Numbers at the top of the bars indicate the numbers of communications per agent in which the lowest SSD is obtained.

Notably, given that both scenarios (1b and 2b) involve consonant substitution in all types of communications, they lead to some amount of change even under the R_{BI} value of 0.0, that is, when the population of agents contains no bilinguals at all. Crucially, however, emergent varieties with the lowest SSD values across all scenarios are generated in populations with bilinguals (see below). In other words, in terms of resemblance to present-day Duoxu, the evolution of the consonant system of Duoxu_PCVG is more consistent with the hypothesis of contact influence from SWM (as reflected in the model by consonant exchange between Duoxu_PCVG and SWM) than with a language-internal development.

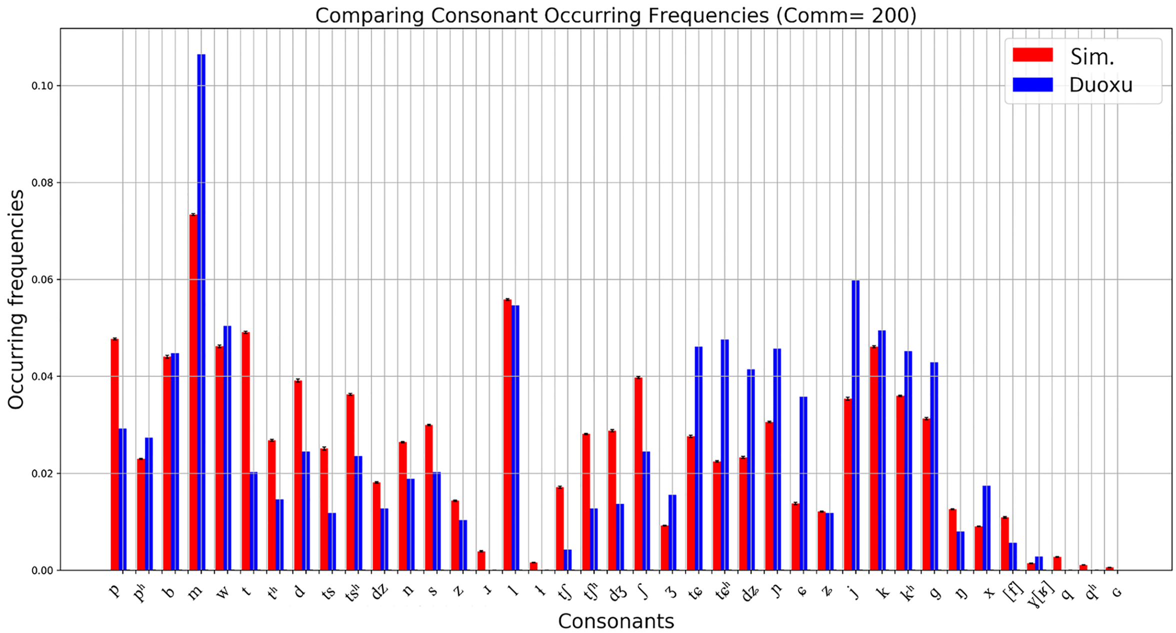
The lowest SSD values in scenarios 1b and 2b are obtained under a small ratio of bilinguals, namely, R_{BI}=0.1 and R_{BI}=0.2, respectively. Figure 5 compares these emergent varieties to Duoxu and Duoxu_PCVG.



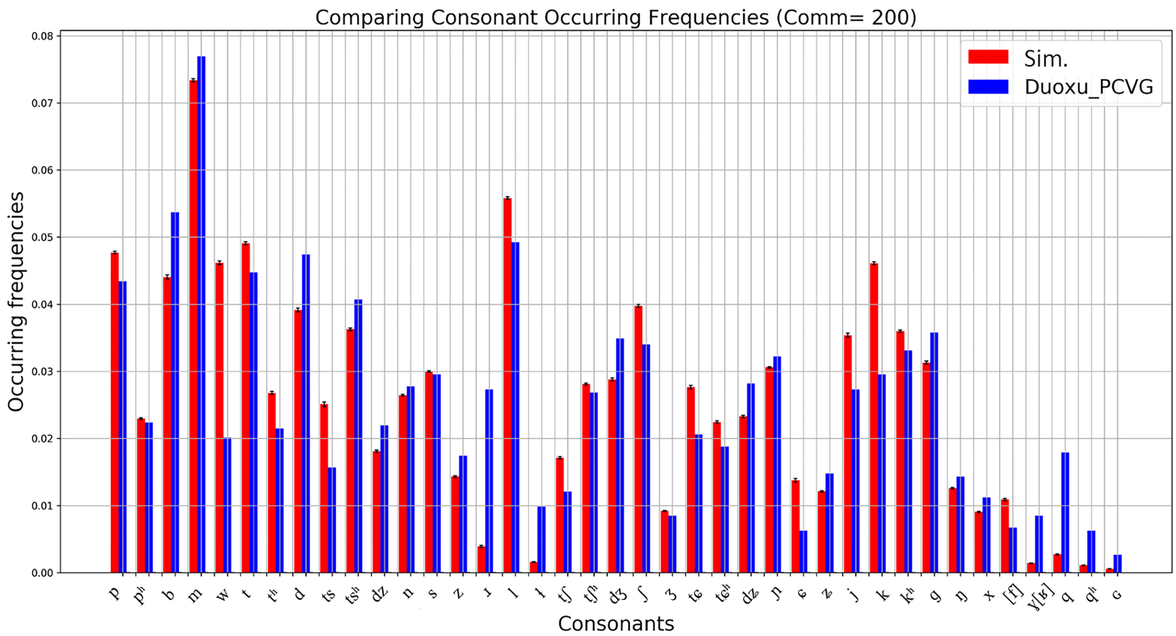
(a)


Comparing Consonant Occurring Frequencies (Comm= 1000)
Sim.
Duoxu_PCVG
Occurring frequencies
Consonants

(b)


Comparing Consonant Occurring Frequencies (Comm= 200)
Sim.
Duoxu
Occurring frequencies
Consonants

(c)



(d)

Figure 5. (a), (b) The emergent variety with the lowest SSD value in scenario 1b (Sim.), as compared to Duoxu (a) and Duoxu_PCVG (b). That emergent variety is obtained with R_{BI}=0.1 and at 10 communications per agent. (c), (d) The emergent variety with the lowest SSD value (Sim.) in scenario 2b, as compared to Duoxu (c) and Duoxu_PCVG (d). That emergent variety is obtained with R_{BI}=0.2 and at 2 communications per agent. The emergent varieties are averaged over 20 runs in the same setting.

In relation to the diagnostic types of changes between Duoxu_PCVG and Duoxu, as outlined in Section 3 – (a) the loss in Duoxu of the consonants /ɬ, ɹ, ɣ, ɢ, q, q^h/ and their substitution by the consonants /l, w, g, k, k^h/, and (b) a related increase in the frequency of occurrence of the Duoxu consonants /l, w, g, k, k^h/) – the emergent variety with the lowest SSD value in scenario 1b yields the best resemblance to Duoxu (Fig. 5a). In this variety, the set of substitution changes is nearly implemented (albeit with some marginal presence of q and a complete loss of ɣ), and at the same time, the compensatory increase in the frequency of occurrence of the consonants /l, w, g, k, k^h/ is comparable with that in Duoxu. By contrast, in the emergent variety with the lowest SSD value in scenario 2b, all consonants to be substituted are still marginally present in the system (Fig. 5c). Therefore, while, in terms of overall consonant frequencies, the emergent variety with the lowest SSD value in scenario 2b provides the best match to Duoxu of all emergent languages with the lowest SSD values, in terms of the implementation of the diagnostic set of changes it is only second to the emergent variety with the lowest SSD value in scenario 1b.

Our statistical results suggest that, when consonant exchange is limited to bilingual-bilingual communications, changes in the system are not different from those in the reference scenario (post-hoc t-test between scenarios 0 and 1a: $t(260.05) = -0.295$, $p = .768$; between scenarios 0 and 2a: $t(275.41) = 2.040$, $p = .042$; and between scenarios 1a and 2a: $t(247.54) = 2.0351$, $p = .0429$). This suggests that implementation of change as limited to bilingual-bilingual communications may not be a realistic modeling of convergence. By contrast, implementation of the set of substitution changes in all types of communications may represent a more accurate mechanism of change as compared to the real-world case study (post-hoc t-test between scenarios 0 and 1b: $t(236.71) = 11.399$, $p < .0001$; between scenarios 1a and 1b: $t(269.83) = 10.612$, $p < .0001$; between scenarios 0 and 2b: $t(195.19) = 19.255$, $p < .0001$; and between scenarios 2a and 2b: $t(185.89) = 18.391$, $p < .0001$).

Let us now turn to the respective roles of the linguistic and social factors in our simulations (post-hoc t-test between scenarios 1b and 2b: $t(252.02) = 9.334$, $p < .0001$). When consonant substitution occurs in all types of communications, both language-internal factors (low frequency and markedness) and language-external factors (social factors) can effectively lead to Duoxu-like emergent varieties under small ratios of bilinguals. While the implementation of social factors is more effective in lowering SSD values, yielding closer resemblance to Duoxu in terms of overall consonant frequencies, the effect of lower frequency and markedness yields greater resemblance to Duoxu in terms of the actual outcome of the set of diagnostic substitution changes. We tentatively take these results to indicate that both factors are of relevance to the process of convergence. (We note at the same time that according to our statistical analysis, markedness is significant only in combination with the factor types of communications). This is in line with the currently prevailing view in contact linguistics that convergence is constrained by multiple internal and external factors.

6 Discussion

In this study, we modeled the process of convergence in the consonant system of Duoxu under the contact influence of SWM. We tested some intuitively plausible assumptions related to the factors assumed or hypothesized to have an impact on phonological convergence and we examined the role of bilinguals in the propagation of change. The following is a summary of our main findings.

Our simulation results confirm the intuitively plausible assumption in convergence studies that the loss of phonological features (such as consonants) in a language that has undergone convergence is correlated with the low frequency of occurrence and markedness of those features in the combined bilingual repertoire. Our simulations also show that social factors (in particular, prestige and amount of "cultural pressure" of the contact language, as reflected in scenario 2b) are highly efficient in leading to Duoxu-like emergent varieties. This corroborates the prevalent view in contact linguistics that social factors are essential in determining both existence and degree of convergence, just like they are – more broadly speaking – in any other language contact phenomena (e.g. Thomason and Kaufman, 1988; Thomason, 2009: 35–39).

In contrast to the intuitively appealing notion that higher numbers of bilinguals promote convergence changes in a population, our simulation results demonstrate that, when taken in combination with linguistic and social factors that enhance convergence (as those discussed in our study), higher numbers of bilinguals in a population fail to yield emergent varieties that resemble Duoxu. Instead, a few bilinguals, with the entire population participating in the propagation of convergence changes, can effectively lead to a Duoxu-like emergent variety (our scenarios 1b and 2b). This being the case, widespread and stable bilingualism may not necessarily be an indispensable prerequisite for convergence. Our simulation results further suggest that the mechanism of propagation of change in situations of convergence (that is, in language contact settings) may not be different from those for language change in monolingual settings, as studied in sociolinguistics (e.g. Croft, 2000; Labov, 2001). Put differently, while the linguistic behavior of bilinguals may be the *source* of convergence changes, the *mechanism* of propagation of these changes in a population group is likely to be identical to that in all other situations of language change and to be driven by social factors (such as prestige). This finding is consistent with a recent quantitative variationist study of the ongoing convergence of the consonant system of Ersu with that of its contact language SWM (Chirkova et al., 2018). Crucially, that study demonstrates that the ongoing phonological convergence between a minority language (Ersu) and a dominant contact language (SWM) manifests itself in a socially stratified way that is consistent with many of the predictions of classic sociolinguistic principles, which were originally stated in relation to monolingual settings.

In sum, the results of our modeling study (that is, the significance of markedness and social factors in the process of convergence, the mechanism of propagation of change, the role of bilinguals) may be useful for systematic exploration of other cases of convergence in various geographically and typologically diverse languages. At the same time, as the model data can only approximate what occurs in an actual case of convergence, the present conclusions are only tentative. Much rests on the further elaboration of the model in order to straighten out the preliminary results. A more nuanced understanding of convergence and of the role of bilinguals therein would require additional consideration of the factors that inhibit convergence, such as awareness of identity and loyalty toward the group associated with the home language. In relation to our Duoxu case study, such additional factors would explain how the consonant system of this language was stabilized after a presumably brief period of convergent changes in the 18th century to remain largely intact until the latest attestations of Duoxu, and before complete phonological convergence in some speakers was precipitated by the ongoing language shift to SWM.

This study was designed as a pilot study for an initial evaluation of the essentially descriptive notion of convergence in contact linguistics and historical linguistics. Naturally, much remains to be done to refine the model and verify that all assumptions are correct. In particular, more work is required to implement the challenging concept of markedness (cf. Baxter et al., 2009: 7) and to refine the model of bilingual language processing (see footnotes 8 and 9) and the model of language representation (see footnote 10). We hope to address these issues in future research.

Studies in language contact and convergence can profit from computational modeling, in the same way many areas of complex systems science (from physics and evolutionary biology to economics) already have. Indeed, a growing body of literature demonstrates the utility of employing computer modeling in investigating language change (e.g. de Boer, 2001; Wichmann, 2008; Gong et al., 2008; Baxter et al., 2009; Gong, 2010; Gong et al., 2010; Blythe and Croft, 2012; Gong et al., 2012; Gong et al., 2013; Wedel et al., 2013; Chirkova and Gong, 2014; Gong et al., 2014; Jansson et al., 2015; Wang and Minett, 2005; Yao and Chang, 2016; Zhang and Gong, 2014). Simulation provides powerful new methods to explore linguistic data. In addition to general advantages of quantitative methods (such as systematicity or applicability of statistical analysis), specific advantages of applying computer modeling in a case study like ours include possibilities of (a) assessing respective roles of individual factors related to convergence, and (b) overcoming problems faced by studies in language change, which are often, by necessity, exclusively based on synchronic evidence, to investigate the origin and course of change. Computer simulations hence lend themselves as a promising tool for investigation of complex cases of language change in contact settings, as they contribute to a more systematic understanding of language change in contact linguistics and historical linguistics.

Acknowledgments

The study is supported in part by the MOE Project of the Center for Linguistics and Applied Linguistics, Guangdong University of Foreign Studies.

Both authors contributed equally to this paper and are listed alphabetically.

References

Andersen, Roger W. 1982. Determining the linguistic attributes of language attrition. In Richard D. Lambert and Barbara Freed (eds.), *The Loss of Language Skills,* 83-118. Rowley, MA: Newbury House.

Antoniou, Mark, Catherine T. Best, Michael D. Tyler, and Christian Kroos. 2011. Inter-language interference in VOT production by L2-dominant bilinguals: Asymmetries in phonetic code-switching. *Journal of Phonetics* 39(4): 558-570.

Baxter, Gareth J., Richard A. Blythe, William Croft, and Alan J. McKane. 2009. Modeling language change: An evaluation of Trudgill's theory of the emergence of New Zealand English. *Language Variation and Change* 21.2: 257-296.

Blythe, Richard A. and William Croft. 2012. S-curves and the mechanisms of propagation in language change. *Language* 88(2): 269-304.

de Boer, Bart. 2001. *The Origins of Vowel Systems.* Oxford University Press, Oxford.

Bybee, Joan. 2010. Markedness: Iconicity, economy and frequency. In Jae Jong Song (ed.), *Handbook of Linguistic Typology,* 131-147. Oxford: Oxford University Press.

Chirkova, Katia. 2014. The Duoxu language and the Ersu-Lizu-Duoxu relationship. *Linguistics of the Tibeto-Burman Area* 37(1): 104-146.

Chirkova, Katia. 2015. A phonological sketch of Duoxu. *Cahiers de Linguistique Asie Orientale* 44(2): 97-121.

Chirkova, Katia. 2017. Lizu (Ersu). In Randy J. LaPolla and Graham Thurgood (eds.), *The Sino-Tibetan Languages* (2nd ed.), 823-839. New York: Routledge.

Chirkova, Katia and Tao Gong. 2014. Simulating vowel chain shift in Xumi. *Lingua* 152: 65-80.

Chirkova, Katia, James N. Stanford, and Dehe Wang. 2018. A long way from New York City: Socially stratified contact-induced phonological convergence in Ganluo Ersu (Sichuan, China). *Language Variation and Change* 30: 109-145.

Clyne, Michael. 2003. *Dynamics of Language Contact: English and Immigrant Languages.* Cambridge Approaches to Language Contact. Cambridge: Cambridge University Press.

Croft, William. 2000. *Explaining Language Change: An Evolutionary Approach.* Harlow, Essex: Longman.

Desmet, Timothy and Wouter Duyck. 2007. Bilingual language processing. *Language and Linguistics Compass* 1(3): 168-194.

Dressler, Wolfgang. 1972. On the phonology of language death. In Paul M. Peranteau (ed.), *Papers from the Eighth Regional Meeting of the Chicago Linguistic Society 1972,* 448-457. Chicago: Chicago Linguistic Society.

Ferber, Jacques. 1998. *Multi-Agent Systems: An Introduction to Distributed Artificial Intelligence.* Harlow, Essex: Addison-Wesley Publishing.

Foley, William. 1986. *The Papuan Languages.* Cambridge: Cambridge University Press.

Gong, Tao. 2010. Exploring the roles of horizontal, vertical, and oblique transmissions in language evolution. *Adaptive Behaviors* 18(3-4): 356-376.

Gong, Tao, James W. Minett, and William S.-Y. Wang. 2008. Exploring social structure effect on language evolution based on a computational model. *Connection Science* 20(2-3): 135-153.

Gong, Tao, James W. Minett, and William S.-Y. Wang. 2010. A simulation study exploring the role of cultural transmission in language evolution. *Connection Science* 22(1): 69-85.

Gong, Tao, Lan Shuai, and Umberto Ansaldo. 2013. Computer simulation of language convergence. In Gang Peng and Feng Shi (eds.), *Eastward Flows the Great River: Festschrift in Honor of Prof. William S.-Y. Wang on His 75th Birthday,* 244-264. Tianjin, China: Nankai University Press.

Gong, Tao, Lan Shuai, Monica Tamariz, and Gerhard Jäger. 2012. Studying language change using Price equation and Pólya-urn dynamics. *PLOS ONE* 7(3): e33171.
Gong, Tao, Lan Shuai, and Menghan Zhang. 2014. Modeling language evolution: Examples and prediction. *Physics of Life Reviews* 11(2): 280-302.
Grosjean, François. 1982. *Life with Two Languages: An Introduction to Bilingualism.* Cambridge, MA: Harvard University Press.
Grosjean, François. 1989. Neurolinguists, beware! The bilingual is not two monolinguals in one person. *Brain and Language* 36: 3-15.
Grosjean, François. 2004. Studying bilinguals: Methodological and conceptual issues. In Tej K. Bhatia and William C. Ritchie (eds.), *The Handbook of Bilingualism,* 32-63. Oxford: Blackwell Publishing.
Guion, Susan. 2003. The vowel systems of Quichua-Spanish bilinguals: Age of acquisition effects on the mutual influence of the first and second languages. *Phonetica* 60(2): 98-128.
Haspelmath, Martin. 2006. Against markedness (and what to replace it with). *Journal of Linguistics* 42: 25-70.
Hock, Hans Henrich. 1991. *Principles of Historical Linguistics.* Second revised and updated edition. Berlin/New York: Mouton de Gruyter.
Huang, Bufan and Wangmu Renzeng. 1991. Lüsuyu [The Lüsu language]. In Qingxia Dai, Bufan Huang, Ailan Fu, Wangmu Renzeng, and Juhuang Liu (eds.), *Zang-Mianyu shiwu zhong / Fifteen Tibeto-Burman Languages,* 132-152. Beijing: Beijing Yanshan Chubanshe.
Huang, Bufan and Weibin Yin. 2012. Duoxuyu gaikuang [An outline of the Duoxu language]. *Han-Zangyu Xuebao* 6: 49-70.
Jansson, Fredrik, Mikael Parkvall, and Pontus Strimling. 2015. Modeling the evolution of creoles. *Language Dynamics and Change* 5(1): 1-51.
Kirby, Simon. 1999. *Function, Selection and Innateness: The Emergence of Language Universals.* Oxford: Oxford University Press.
Kouwenberg, Silvia. 2001. Convergence and explanations in creole genesis. In Norval Smith and Tonjes Veenstra (eds.), *Creolization and Contact,* 219-247. Amsterdam/Philadelphia: John Benjamins Publishing Company.
Kroll, Judith F., Jason Gullifer, Rhonda McClain, Eleonora Rossi, and María Cruz Martín. 2015. Selection and control in bilingual comprehension and production. In John W. Schwieter (ed.), *The Cambridge Handbook of Bilingual Processing,* 485-507. Cambridge: Cambridge University Press.
Labov, William. 2001. *Principles of Linguistic Change (Vol. 2): Social Factors.* Oxford: Blackwell.
Li, Lan. 2010. Sichuan Muli xian fangyan jilue [The Chinese dialect of Muli, Sichuan: A survey]. *Fangyan* [Dialects] 2: 114-133.
Matras, Yaron. 2009. *Language Contact.* Cambridge: Cambridge University Press.
Muysken, Pieter. 2010. Scenarios for language contact. In Raymond Hickey (ed.), *The Handbook of Language Contact,* 265-281. Singapore: Wiley-Blackwell.
Nishida, Tatsuo. 1973. *Tosu yakugo no kenkyū: shin gengo tosugo no kōzō to keitō* [A study of the Tosu-Chinese vocabulary, Tosu i-yu: the structure and lineage of Tosu, a new language]. Kyoto: Shokado.
Norman, Jerry. 1999 [1st ed. 1988]. *Chinese.* Cambridge Language Surveys. Cambridge: Cambridge University Press.
R Core Team. 2016. R: A language and environment for statistical computing. Vienna: R Foundation for Statistical Computing. URL: https://www.R-project.org/.

Rabus, Achim. 2014. Siblings in contact: The interaction of Church Slavonic and Russian. In Juliane Besters-Dilger, Cynthia Dermarkar, Stefan Pfänder, and Achim Rabus (eds.), *Congruence in Contact-Induced Language Change: Language Families, Typological Resemblance, and Perceived Similarity,* 337-410. Berlin: De Gruyter.

Shepard, Carolyn A., Howard Giles, and Beth A. Le Poire. 2001. Communication accommodation theory. In W. Peter Robinson and Howard Giles (eds.), *The New Handbook of Language and Social Psychology,* 33-78. Chichester: John Wiley & Sons.

Sichuan Sheng Mianning Xian Difangzhi Bianzuan Weiyuanhui 四川省冕宁县地方志编纂委员会 [Mianning County, Sichuan Province, Gazetteer Editing Committee]. 2009. *Mianning Xian Zhi* 《冕宁县志》. Chengdu: Xinan Jiaotong Daxue Chubanshe.

Smith, Kenny and James R. Hurford. 2003. Language evolution in populations: Extending the Iterated Learning Model. In Wolfgang Banzhaf, Thomas Christaller, Peter Dittrich, Jan T. Kim, and Jens Ziegler (eds.), *Proceedings of the 7th European Conference on Artificial Life,* 507-516. Berlin: Springer-Verlag.

Thomason, Sarah Grey. 2009. Contact explanations in linguistics. In Raymond Hikey (ed.), *The Handbook of Language Contact,* 31-47. Malden, MA/Oxford: Blackwell Publishers.

Thomason, Sarah Grey and Terrence Kaufman. 1988. *Language Contact, Creolization, and Genetic Linguistics.* Berkeley/Los Angeles/Oxford: University of California Press.

Steels, Luc. 1997. The synthetic modeling of language origins. *Evolution of Communication* 1(1): 1-34.

Sun, Hongkai. 1983. Liujiang liuyu de minzu yuyan ji qi xishu fenlei [Minority languages of the Six River Valley and their genetic classification]. *Minzu Xuebao* [Scholarly Journal of Nationalities] 3: 99-274.

Wang, Dehe. 2010. *Ersu zangzu wenhua yanjiu* [Study of Ersu Tibetan culture]. Chengdu: Sichuan University Press.

Wang, William S.-Y. 1969. Competing sound changes as a course of residue. *Language* 45: 9-25.

Wang, William S.-Y. and James W. Minett. 2005. The invasion of language: Emergence, change, and death. *Trends in Ecology and Evolution* 20(5): 263-269.

Wedel, Andrew, Scott Jackson, and Abby Kaplan. 2013. High functional load inhibits phonological contrast loss: A corpus study. *Cognition* 128(2): 179-86.

Weinreich, Uriel. 1974 [1st ed. 1953]. *Languages in Contact: Findings and Problems.* The Hague/Paris: Mouton.

Wichmann, Søren. 2008. The emerging field of language dynamics. *Language and Linguistics Compass* 2(3): 442-455.

Winford, Donald. 2003. *An Introduction to Contact Linguistics.* Malden, MA: Blackwell Publishers.

Yao, Yao and Charles B. Chang. 2016. On the cognitive basis of contact-induced sound change: Vowel merger reversal in Shanghainese. *Language* 92(2): 433-467.

Yuan, Jiahua. 2001 [1st ed. 1960]. *Hanyu fangyan gaiyao* [Outline of Chinese dialects]. Beijing: Yuwen Chubanshe.

Yu, Dominic. 2012. *Proto-Ersuic.* PhD dissertation, University of California, Berkeley.

Zhang, Menghan and Tao Gong. 2014. Principles of parametric estimation in modeling language competition. *Proceedings of the National Academy of Sciences of the U.S.A.* 110(24): 9698-9703.

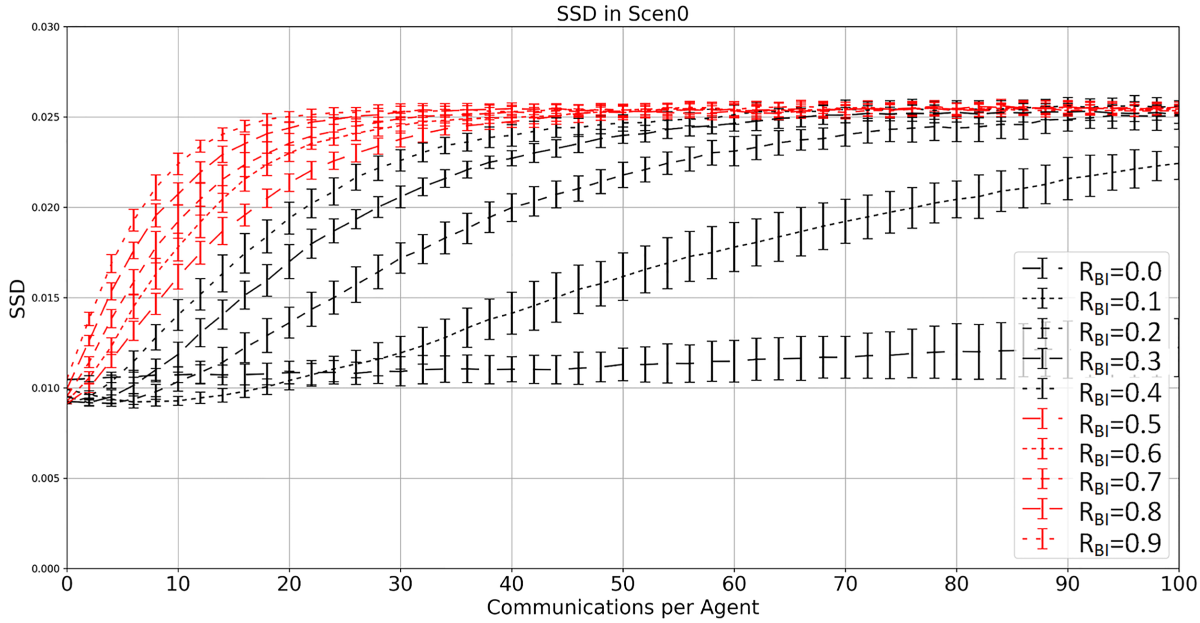
Supplementary Materials

Consonant frequencies in SWM, Lizu (Duoxu_PCVG), and Duoxu

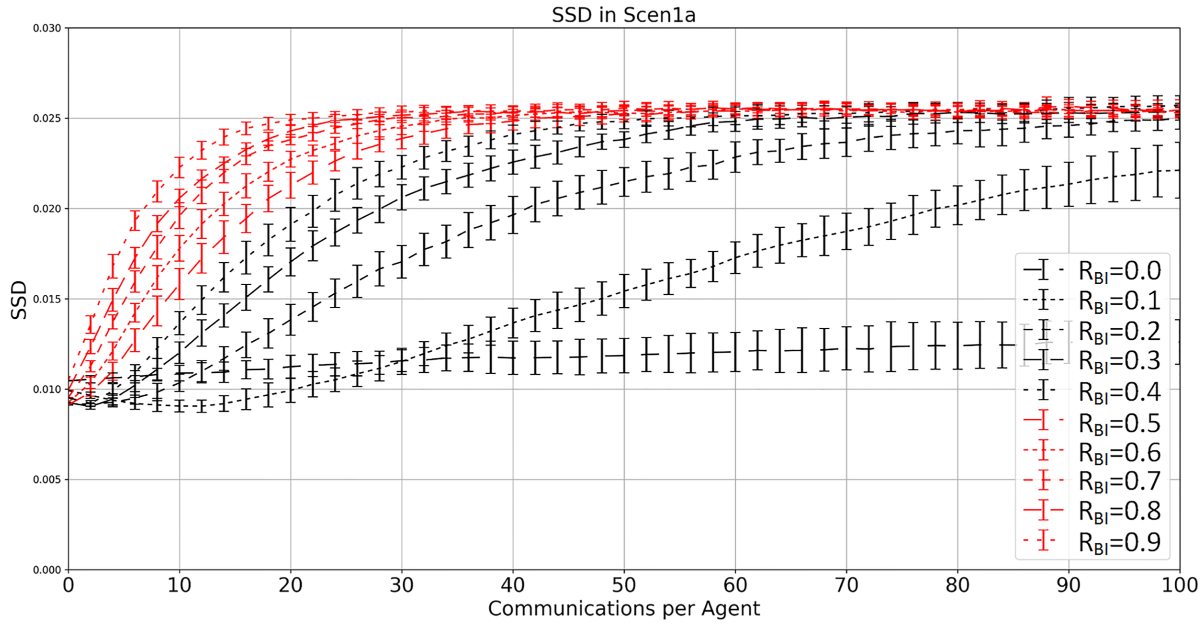
Table A1. Consonant frequencies in SWM, Lizu, and Duoxu ("Cons."=consonant, "Occur."=occurrence, "Freq."=frequency, "Diff."=difference). Empty cells indicate that the language does not have those consonants. Frequency columns show the normalized frequency of a consonant calculated from its number of occurrence. Difference columns show the occurrence and frequency differences of corresponding consonants between Lizu and Duoxu: "+" means that Duoxu has more occurrences (or higher frequency) of a consonant, "-" means that Duoxu has fewer occurrences (or lower frequency) of a consonant.

Tag	Cons.	SWM		Lizu		Duoxu		Dif.	Dif.
		Occur.	Freq.	Occur.	Freq.	Occur.	Freq.	in occur.	in freq.
1	p	140	0.0673	97*	0.0434	62	0.0292	-35	-0.3608
2	pʰ	55	0.0265	50*	0.0224	58	0.0273	+8	+0.1600
3	b			120*	0.0537	95	0.0447	-25	-0.2083
4	m	116	0.0558	165*	0.0738	226*	0.1065	+61	0.3697
5	w	55	0.0265	45	0.0201	107	0.0504	+62	1.3778
6	t	149	0.0717	100	0.0447	43	0.0203	-57	-0.5700
7	tʰ	111	0.0534	48*	0.0215	31	0.0146	-17	-0.3542
8	d			106*	0.0474	52*	0.0245	-54	-0.5094
9	ts	148	0.0712	35	0.0157	25	0.0118	-10	-0.2857
10	tsʰ	30	0.0144	91*	0.0407	50	0.0236	-41	-0.4506
11	dz			49*	0.0219	27	0.0127	-22	-0.4490
12	n	43	0.0207	54	0.0242	40	0.0188	-14	-0.2593
13	s	62	0.0298	66	0.0295	43	0.0203	-23	-0.3485
14	z			39	0.0175	22	0.0104	-17	-0.4359
15	ɹ			61	0.0273			-61	-1.0000
16	l	100	0.0481	110	0.0492	116	0.0546	+6	0.0546
17	ɬ			22	0.0098			-22	-1.0000
18	tʃ	82	0.0394	27	0.0121	9	0.0042	-18	-0.6667
19	tʃʰ	70	0.0337	60*	0.0268	27	0.0127	-33	-0.5500
20	dʒ			78*	0.0349	29	0.0137	-49	-0.6282
21	ʃ	140	0.0673	76*	0.0340	52	0.0245	-24	-0.3158
22	ʒ	24	0.0115	19*	0.0085	33	0.0155	+14	0.7368
23	tɕ	127	0.0611	46	0.0206	98	0.0462	+52	1.1304
24	tɕʰ	84	0.0404	42*	0.0188	101	0.0476	+59	1.4048
25	dʑ			63*	0.0282	88	0.0415	+25	0.3968
26	ɲ	48	0.0231	65	0.0291	97	0.0457	+32	0.4923
27	ɕ	103	0.0495	14	0.0063	76	0.0358	+62	4.4285
28	ʑ			33	0.0148	25	0.0118	-8	-0.2424
29	j	152	0.0731	61	0.0273	127	0.0598	+66	1.0820
30	k	114	0.0548	66	0.0295	105	0.0495	+39	0.5909
31	kʰ	51	0.0245	74*	0.0331	96	0.0452	+22	0.2973
32	g			80*	0.0358	91	0.0429	+11	0.1375
33	ŋ	9	0.0043	24	0.0107	17	0.0080	-7	-0.2917
34	x			25*	0.0112	37	0.0174	+12	0.4800
35	[f]	66	0.0317	15	0.0067	12	0.0057	-3	-0.2000
36	ɣ [ʁ]			19	0.0085	6	0.0028	-13	-0.6842
37	q			40	0.0179			-40	-1.0000
38	qʰ			14	0.0063			-14	-1.0000
39	ɢ			6	0.0027			-6	-1.0000
40	h̃ [ɦ̃]			30	0.0134			-30	-1.0000

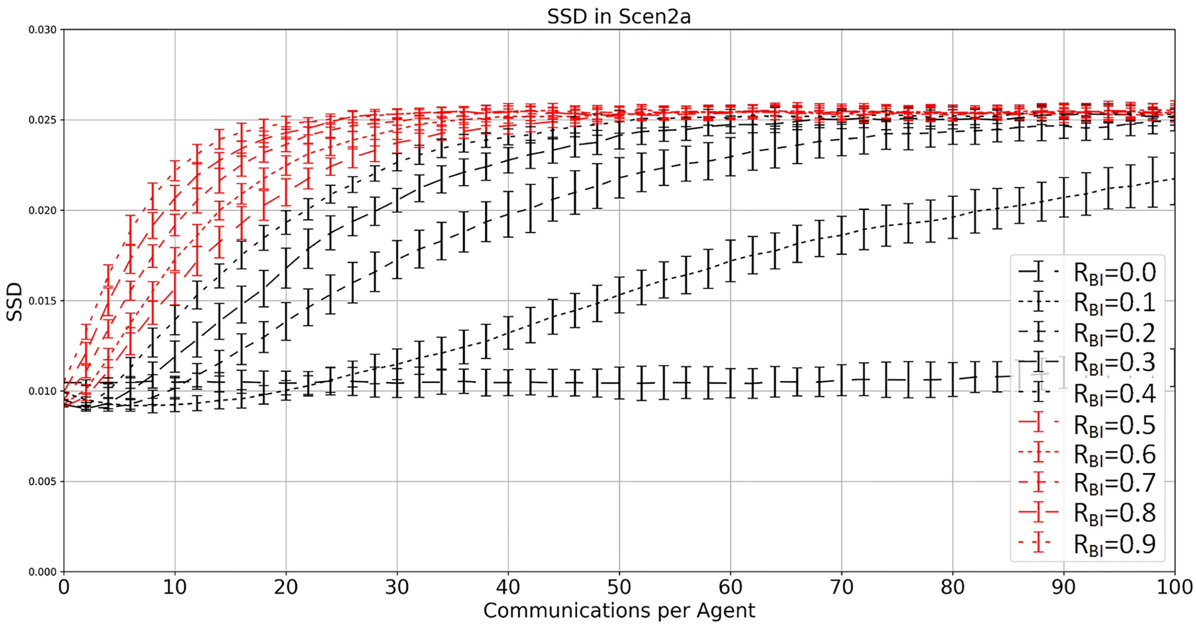
*: also includes occurrences of these initial consonants in clusters (e.g. /bɹ, bz, Nd, Ntɕ^h/ etc.)



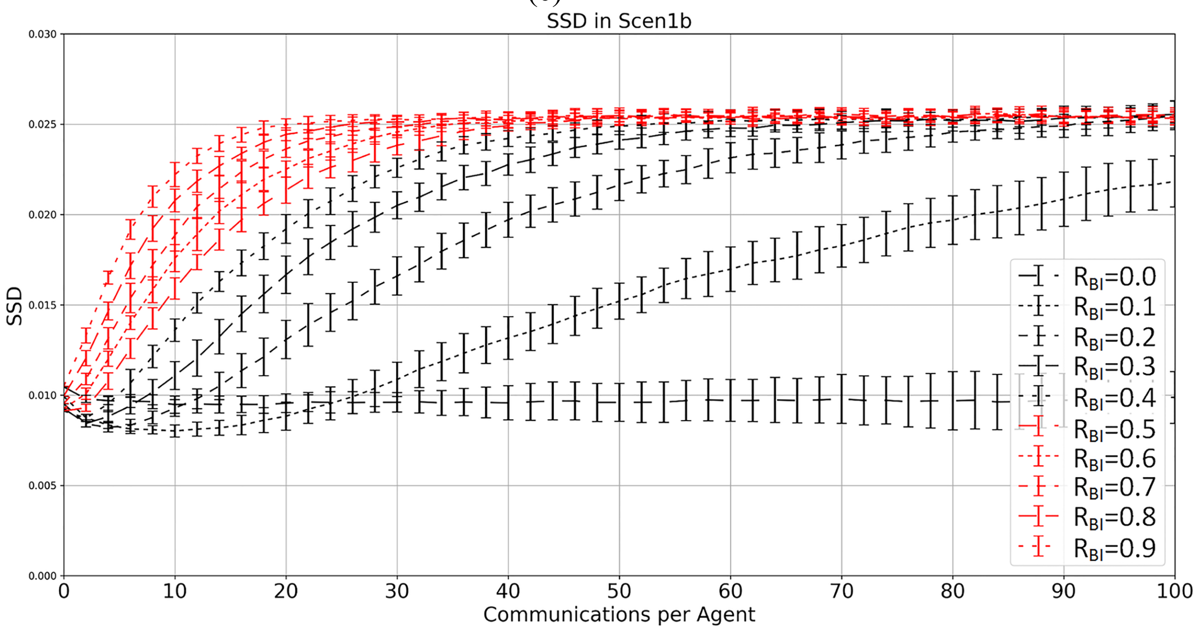
(a)



(b)



(c)



(d)

Figure A1. Mean SSD values obtained throughout 10,000 communications in scenario 0 (a), 1a (b), 2a (c), and 1b (d). Each curve in each figure corresponds to the result under a particular R_{BI}. Results are averaged over 20 runs in each condition.

Effects of F_{MKD} and F_{ADJ}

In this section, we briefly discuss the effects of the parameters F_{MKD} and F_{ADJ}. For the sake of simplicity, we fix the values of all other model parameters and vary the values of F_{MKD} and F_{ADJ} to examine their effects on the dynamics of the system. The results are based on the average SSD values obtained at each sampling point over 20 runs of the same setting. Since most discussion in the main text concerns a small R_{BI}, here, we fix R_{BI} as 0.1. The values of the other parameters (such as N_P, N_C, and N_{CONS}) are the same as those in the main text. For each of the two parameters F_{MKD} and F_{ADJ}, we select two additional values, one larger and the other smaller than the value set in the simulations in the main text.

F_{MKD} is used in scenarios 1a and 1b. Together with the current value of 0.01 used in the simulations, we select another two values (0.005 and 0.05) for comparison. Figure A2 shows the average SSD values throughout the simulations in scenario 1a under these values (0.005, 0.01, and 0.05). Note that in this scenario, where low frequency of occurrence and markedness take effect only in bilingual-bilingual communications, the SSD values under all three F_{MKD} values are similar throughout the simulations.

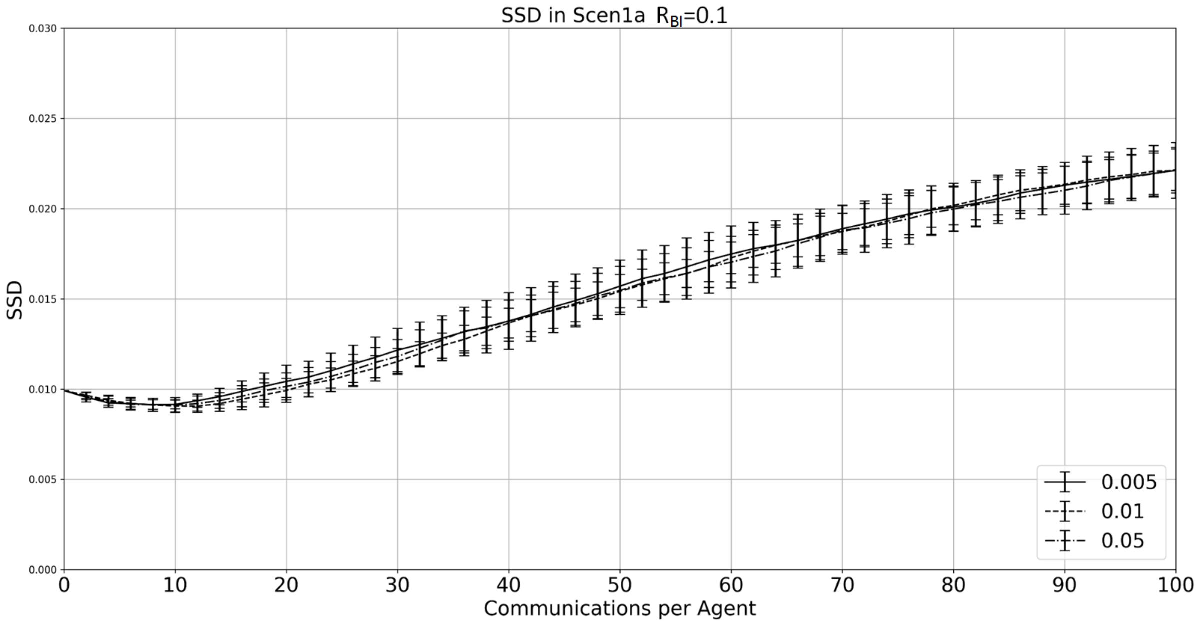


Figure A2. Mean SSD values obtained throughout 10,000 communications in scenario 1a under R_{BI}=0.1. Each curve represents the average results over 20 runs under a particular F_{MKD}.

F_{ADJ} takes effect in all four scenarios. Here, we only consider scenario 2b, and together with the current value (0.002) in the simulations reported in the main text, we choose another two values (0.001 and 0.005). Figure A3 shows the average SSD values throughout the simulations in the two scenarios under the three values of F_{ADJ} (0.001, 0.002, and 0.005). It is shown that in this scenario, the SSD values under all three F_{ADJ} values are similar throughout the simulations.

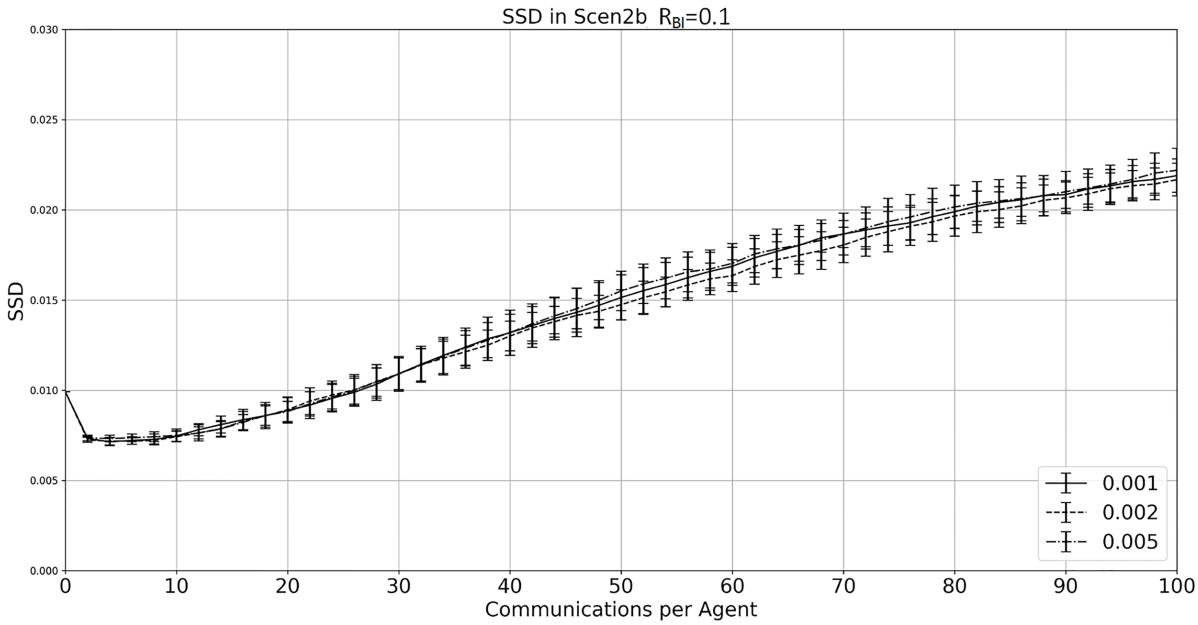


Figure A3. Mean SSD values obtained throughout 10,000 communications in scenario 2b under R_{BI}=0.1. Each curve represents the average results over 20 runs under a particular F_{ADJ}.