COGNITIVE SCIENCE
A Multidisciplinary Journal

Check for updates

Cognitive Science 48 (2024) e70011

ISSN: 1551-6709 online
DOI: 10.1111/cogs.70011

Adapting to Individual Differences: An Experimental Study of Language Evolution in Heterogeneous Populations

Mathilde Josserand,[a,b] François Pellegrino,[a] Oxana Grosseck,[c,d] Dan Dediu,[e,f,g] Limor Raviv[c,h]

[a]*Laboratoire Dynamique du Langage, Université Lyon 2/CNRS UMR 5596*
[b]*Laboratoire Eco-Anthropologie UMR 7206, CNRS/MNHN/Université Paris Cité*
[c]*LEADS Group, Max Planck Institute for Psycholinguistics*
[d]*Donders Institute for Brain, Cognition and Behaviour, Radboud University Nijmegen*
[e]*Department of Catalan Philology and General Linguistics, University of Barcelona*
[f]*University of Barcelona Institute for Complex Systems (UBICS)*
[g]*Catalan Institute for Research and Advanced Studies (ICREA)*
[h]*cSCAN, University of Glasgow*

Received 2 July 2024; received in revised form 18 September 2024; accepted 18 October 2024

Abstract

Variations in language abilities, use, and production style are ubiquitous within any given population. While research on language evolution has traditionally overlooked the potential importance of such individual differences, these can have an important impact on the trajectory of language evolution and ongoing change. To address this gap, we use a group communication game for studying this mechanism in the lab, in which micro-societies of interacting participants develop and use artificial languages to successfully communicate with each other. Importantly, one participant in the group is assigned a keyboard with a limited inventory of letters (simulating a speech impairment that individuals may encounter in real life), forcing them to communicate differently than the rest. We test how languages evolve in such heterogeneous groups and whether they adapt to accommodate the unique characteristics of individuals with language idiosyncrasies. Our results suggest that language evolves differently in groups where some individuals have distinct language abilities, eliciting more innovative elements at the cost of reduced communicative success and convergence. Furthermore, we observed strong partner-specific accommodation to the minority individual, which carried over to the group level. Importantly, the degree of group-wide adaptation was not uniform and depended on participants' attachment to established language forms. Our findings provide compelling evidence

Correspondence should be sent to Mathilde Josserand, Laboratoire Dynamique du Langage, Université Lyon 2 - CNRS UMR 5596, Lyon, 69007, France. E-mail: mathilde.josserand@gmail.com

that individual differences can permeate and accumulate within a linguistic community, ultimately driving changes in languages over time. They also underscore the importance of integrating individual differences into future research on language evolution.

Keywords: Language evolution; Individual differences; Language change; Micro-societies; Variability; Linguistic alignment

1. Introduction

Languages demonstrate a remarkable capacity to adapt to their environment (Everett, Blasi, & Roberts, 2015; Nölle, Fusaroli, Mills, & Tylén, 2020; Regier, Carstensen, & Kemp, 2016), their culture (MacKeigan & Muth, 2006; Majid & Kruspe, 2018), and even the type of network (Lupyan & Dale, 2010; Nettle, 2012) in which they are embedded. Recent studies have also revealed intriguing evidence of languages adapting to the anatomical and physiological biases of their speakers (Dediu, Janssen, & Moisik, 2019; Moisik & Dediu, 2017). For example, the distinctive bite configuration observed in hunter-gatherer populations—which is induced by an intense tooth wear diet—renders the production of labiodental sounds slightly more challenging compared to the typical bite configuration of agriculturalists, which may contribute to the lower occurrence of labiodental sounds in their languages' sound repertoire (Blasi et al., 2019). These studies suggest the existence of different, usually very small, patterns of variation in human anatomy and physiology across people from various ethnolinguistic groups. These differences lead to subtle variations in the produced language, which may give rise to cross-linguistic differences in the long-term (Dediu, Janssen, & Moisik, 2017). Indeed, as language is repeatedly used and transmitted within a population over generations, even weak biases could be amplified, thus playing a role in shaping the landscape of linguistic diversity (Kirby, Dowman, & Griffiths, 2007).

This literature emphasizes the rich pool of diversity present among different populations, but also sheds light on the substantial variation in anatomico-physiological features that exists between the speakers of the same language. Here, we shift our focus from between-group differences to specifically examine within-population variation. This within-population in anatomy and physiology is just one of many sources of language variation; other differences in linguistic capacity (Labov, 1979) or cognitive ability (Kapnoula, 2016; Kong & Edwards, 2016; Lev-Ari & Peperkamp, 2013, 2014) may also affect language production and perception. Theories of sociolinguistics emphasize additional contributing factors such as gender, age, socioeconomic background, and regional exposure, all of which contribute to the intricate patterns of linguistic diversity within languages (Byrd, 1992; Foulkes & Docherty, 2006; Labov, 1994, 2001). This variation leads to significant individual differences in crucial language aspects, with sound production being one notable example (Allen, Miller, & DeSteno, 2003; Johnson, Ladefoged, & Lindau, 1993; Mielke, Baker, & Archangeli, 2016; R. Newman, 1997; R. S. Newman, Clouse, & Burnham, 2001; Peterson & Barney, 1952; Zellou, 2017). For example, in Dutch, /r/ can be realized among others as a tap [ɾ], approximant [ɹ], trill [r] or [ʀ] or even as a fricative [ʁ] (Sebregts, 2014). As such, variation in language is ubiquitous, spanning across all levels within a population, encompassing differences not only between social groups but also within them (Fillmore, Kempler, & Wang, 2014; Yu & Zellou, 2019). In this study, we aim to investigate whether individual variation in language present within a group, such as the inability to produce a sound, might influence the language employed by the entire group. To do so, we use an experimental approach involving groups of four participants communicating using a miniature artificial language, with one participant assigned a biased keyboard with a limited inventory of letters.

Remarkably, individuals may naturally embrace the language variations of others. Research shows that individuals have an unconscious inclination to adjust their language during communication, mirroring that of their conversation partner (also known as *linguistic alignment*[1]). This interactive alignment, which leads to the synchronization of linguistic and mental representations between speakers, is key to making conversation effortless and efficient (Garrod & Pickering, 2004). This alignment can occur for a range of linguistic parameters, such as phonetics (Pardo, 2006), speaking rate (Giles, Coupland, & Coupland, 1991), word choices (e.g., Brennan & Clark, 1996), grammatical structure (Branigan, Pickering, Pearson, & McLean, 2010; Garrod & Pickering, 2004), and even nonverbal behaviors like gestures and facial expressions (Chartrand & Bargh, 1999). Numerous explanations for these adjustments have been proposed: it could be a result of priming (Branigan et al., 2010), a way to increase the chances of successful communication (e.g., Bortfeld & Brennan, 1997; Clark, 1996; Garrod & Pickering, 2004), or a strategy to increase social bonding (van Baaren, Holland, Steenaert, & van Knippenberg, 2003). Some individuals demonstrate a greater inclination toward linguistic alignment than others. These individual differences in linguistic alignment are associated with the personal characteristics of both the individuals themselves and their communication partners. These personal characteristics encompass hierarchical position in the workplace, leadership roles, gender, ethnic origin, likability, social desirability, and position within a social network (Bilous & Krauss, 1988; Gnisci, 2005; Jones, Gallois, Callan, & Barker, 1999; Jones, Cotterill, Dedney, Muir, & Joinson, 2014; Natale, 1975; Noble & Fernandez, 2015; Willemyns, Gallois, Callan, & Pittam, 1997).

Thus, we all have our unique way of expressing language and we adapt differently to the speakers with whom we are communicating. While this alignment mostly concerns pair interaction, the substantial individual differences in language raise an important question at the group level: How may these individual differences shape the trajectory of language evolution?

Studies of language evolution typically overlook these individual differences, focusing instead on universal cognitive and communicative biases. For example, some research has identified universal pressures for languages, particularly in semantic categories, to be both simple and informative (Carr, Smith, Culbertson, & Kirby, 2020). This emphasis on universal biases is motivated by the intention to offer insights into the core aspects of language that are shared by all human languages, to better understand the origins, development, and constraints of language. This approach has also led to an increased emphasis on group-level patterns and phenomena typical of populations (Yu & Zellou, 2019), aimed at understanding how languages are shaped by environmental and cultural factors (De Busser, 2015; Raviv, Meyer, & Lev-Ari, 2019b). The use of statistical analyses that focus on averaging data and excluding outliers may also contribute to this oversight. Yet, treating groups as homogeneous units can lead to a substantial loss of valuable information for understanding language change (Yu & Zellou, 2019).

Notably, while languages adapt to universal cognitive and communicative biases (such as biases in learning; Culbertson, 2012), they could also accommodate the biases of only a subset of their speakers. For example, The Al-Sayyid Bedouin community in Israel and the Kata-Kolok community in Indonesia experience 2−5% of their population being born with congenital deafness (compared to approximately 0.1% in the United States; De Vos, 2011; Fox, 2008). In response to this rare situation where there is a relatively high degree of deafness in the population, both communities have developed local sign languages that are used by both deaf and hearing members. These sign languages showcase a sophisticated structure and appear to have evolved spontaneously over generations through interactions between deaf and hearing members (Sandler, Meir, Padden, & Aronoff, 2005). Without considering this minority's influence, our grasp of how these languages have emerged and evolved would remain incomplete. Furthermore, linguistic alignment between users can lead to generalized changes. For example, linguistic alignment between adult native and non-native speakers can lead to the simplification of the language (Frank & Smith, 2018). Moreover, using an artificial language learning task, Fehér, Wonnacott, and Smith (2016) found that participants aligned with their conversational partner (i.e., by adopting the partner's preferred word order) and that interaction led to a tendency toward language regularization, as indicated by a reduction in variability and an increase in the prevalence of the dominant word order. Adding to these findings, Féher et al. (2019) showed that linguistic accommodation is asymmetrical, such that depending on users' experience with the language (and specifically, the level of variability within it) they may be more or less prone to adapting to their interlocutor, leading to systematic changes in the language toward decreased variability. Even though these examples concern individual differences that manifest in comprehension, these examples highlight the idea that small-scale interactions may influence population-level linguistic phenomena, and that it is, therefore, crucial to consider individual differences when studying the evolution of language.

However, little is known about the exact dynamics underlying the processes of language adaptation to a minority, let alone in the context of the evolution of language. Recently, Josserand, Allassonnière-Tang, Pellegrino, and Dediu (2021) tackled this issue by using agent-based models in complex communicative networks. In this model, Bayesian agents interact using an abstract language feature. While some agents lack prior expectations about language, others are inherently predisposed toward a specific linguistic variant, and are referred to as *biased agents*. This model allowed the investigation of several parameters, such as the percentage of biased individuals or the bias strength necessary to influence the language of a population. Contrary to the intuitive view suggesting that the biased minority would eventually adopt the language of the unbiased majority, this research suggested that the emergent shared language of the entire group is, in fact, influenced by the biases held by some individuals, even when they are the minority. Additionally, the spread of individual differences depended on the local position of individuals in the network: when biased agents strategically hold many connections or act as bridges between multiple agents, their impact on the group's language becomes more pronounced (Fagyal, Swarup, Escobar, Gasser, & Lakkaraju, 2010; Josserand, Allassonnière-Tang, Pellegrino, Dediu, & de Boer, 2024). These results also resonate with the work of Navarro, Perfors, Kary, Brown, and Donkin (2018), who showed that in heterogeneous populations, languages transmitted via iterated learning (i.e., chains of multiple generations of participants where the output of one person becomes the input for the next person in the chain) can become unpredictable and often distorted by the Bayesian learners with the strongest beliefs toward a specific rule. While such agent-based models are valuable for exploring parameters relative to network structure, they may not accurately represent the dynamics of adaptation at play at the individual level. Specifically, in the two models mentioned earlier (Josserand et al., 2021; Navarro et al., 2018), communication is Bayesian (Griffiths & Kalish, 2007), which could arguably diverge from real-world communication dynamics (Ferdinand & Zuidema, 2009) or not be fully tractable (Woensdregt et al., 2021). Thus, it is essential to complement these computational models with laboratory-controlled experiments, which allow us to observe the live formation of languages with real interacting human participants.

Laboratory-controlled experiments provide robust findings regarding the processes at play in language evolution and explore causal relationships. They often involve the laboratory-based learning of artificial languages and the observation of their evolution during transmission among participants. These approaches have been used to simulate the emergence of systematicity and compositionality in miniature languages, as well as examine the trajectory of language change over generations (Kirby, Cornish, & Smith, 2008, 2014; Raviv, Meyer, & Lev-Ari, 2019a; Scott-Phillips & Kirby, 2010; Tamariz & Kirby, 2016). A new paradigm entails embedding participants within micro-societies, where they collaboratively create languages from scratch as they engage in successive interactions involving both familiar and novel referents. This approach is especially valuable in understanding how individuals adapt to one another's languages within the context of language evolution. Moreover, since it enables the manipulation of group composition, this approach has also shed light on how the social environment shapes languages. Notably, research has revealed that language evolves differently within communities of varying sizes (Raviv et al., 2019b) and structures (Raviv, Meyer, & Lev-Ari, 2020). Since this micro-society communication paradigm makes it possible to artificially introduce variation within a group in a controlled manner, it serves as the ideal framework for observing the influence of individual differences on language evolution and to study how this may impact the group's language.

In the current study, we used this experimental approach to test how language evolves in heterogeneous populations of interacting participants. Our experiment involves small groups of four individuals who attempt to communicate successively in alternating pairs using an artificial language. In each group, one participant is given a biased keyboard restricting their ability to produce specific letters, thus simulating a speech variation or speech impairment that individuals may experience in their real lives. Our aim is to understand how this biased participant affects the language of the group. Will the majority wipe away the biases of the minority and use a language featuring all the available letters, or conversely, will the collective language of the group adapt to accommodate the biases of the minority? Furthermore, our study allows us to investigate how the personal characteristics of the participants might promote or hinder this adaptive process. Specifically, we explore how the participants' prosociality and cognitive flexibility may affect their tendency to adjust to others. Prosocial behavior encompasses dimensions of empathy and perspective-taking (Caprara, Steca, Zelli, & Capanna, 2005), which entail feeling others' needs in order to react adequately to them. As individuals who possess high perspective-taking abilities often mirror the actions of those they interact with (Chartrand & Bargh, 1999), we expect that prosocial individuals would be more inclined to adapt to the biases of their interlocutor. On the other hand, being more cognitively flexible (i.e., having the ability to adapt flexibly to a constantly changing environment; Cools, 2015) is, among others, associated with a better theory-of-mind in language (Jacques & Zelazo, 2005), and by definition, may also be related to how easily participants can adapt their language to a changing environment. This prompts us to consider that cognitive flexibility could predict the extent to which participants will change their language to align to others. With this experimental approach, our study aims to shed light on the complex relationship between individual differences and language evolution, and the processes at play during linguistic adaptation.

In this study, we used an experimental method adapted from the group communication game (Raviv et al., 2019a, 2019b), where small groups of four participants must communicate using a miniature artificial language. After being exposed to an initial set of labels, participants engaged in face-to-face interactions in alternating dyads over several rounds. During each round, one participant described an image using a label, while their partner needed to select the correct image from a set of distractors. To introduce individual differences, one participant (the "biased participant") was restricted from using two out of eight letters on the keyboard, simulating a speech variation or impairment. We asked whether the group's collective language will adapt to accommodate the biased participant (i.e., by reducing the use of the "unavailable" letters or avoiding them altogether), or whether the majority will dominate—resulting in a language that includes all letters. To preview our main findings, our results show that individuals adapted their languages to fit the variations introduced by the biased individuals, and that such local adaptations often percolated to the rest of the community and accumulated over time—causing languages to change from their initial state to better accommodate the minority.

2. Methods

Our experiment is a group communication game in which four participants, in a fully connected network, interact face-to-face using a computer interface. After an initial learning phase of an artificial miniature language, participants interact in alternating dyads over multiple rounds, playing a referential game. In each communication round, one participant produces a label to describe an image to their partner, who must then identify the correct item from a set of alternatives displayed on the screen. We introduce a production (articulation-like) bias by restricting a designated participant's keyboard input, preventing them from using two specific letters. This individual is referred to as the "*biased*" participant, while the others are called "*unbiased*" participants.

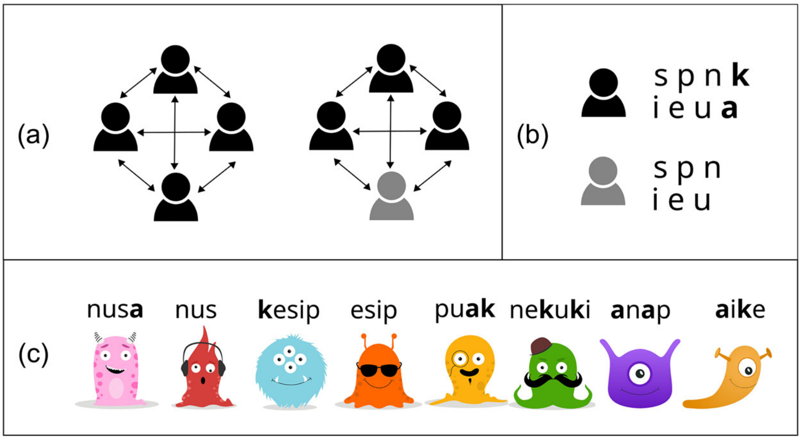

Illustration 1. Panel A: We tested two types of groups: control homogeneous (without biased participants, on the left) and heterogeneous (with one biased participant, in gray, on the right). Panel B: Letter inventories for unbiased (top) and biased (bottom) participants. Panel C: The form-meaning mappings in the miniature language: eight alien images, each associated with a unique label. Biased letters are shown in bold.

2.1. Participants

We tested 56 participants between the ages of 18 and 37 ($M = 24$, $SD = 5$; 38 women), for a total of 14 groups (seven heterogenous and seven control groups). All participants were proficient Dutch speakers.[2] The study was conducted at the Max Planck Institute for Psycholinguistics (Nijmegen, Netherlands), and participants received a compensation of 21 euros for their participation. An ethical certificate was obtained for this study (certificate number: ECSW-LT-2023-2-10-71121, Social Sciences Ethics Committee, the Netherlands). Participants gave their informed consent to participate.

2.2. Production bias

Participants communicate exclusively by typing on their keyboard, with any other form of communication (e.g., speaking, pointing, gesturing) being strictly prohibited. They are explicitly instructed not to use Dutch, English, or any other known languages when typing labels, and to avoid abbreviations or any familiar symbols. Unbiased participants can type the following letters: consonants “s,” “n,” “p,” “k” and vowels “e,” “i,” “u,” “a.” In contrast, the biased participant lacks the ability to produce the “k” and “a” letters, which we will refer to as the *biased letters* throughout the paper (see Illustration 1, Panel B). As the experiment began, the biased participant became aware of their inability to produce certain letters, but remained ignorant both to the purpose of the experiment and to the fact that they were the only biased participant of the group. Importantly, the unbiased participants were not informed that one participant may have a production bias. To avoid any interfering bias based on age or gender, we ensured that any participant in the minority within their group (e.g., one woman

Illustration 2. Visual representation of the procedure. See Section Procedure for more information.

in association with three men) was not placed as the biased participant. Except for that, we randomly selected which participant would be assigned as the biased participant.

To confirm the lack of explicit awareness of our manipulation, at the end of the experiment, we collected participants' subjective feedback by inquiring about their experience. Indeed, none of the participants realized that there was one individual in the group who was systematically different. When asked why they thought communication was challenging sometimes, participants attributed it to others having a significantly poorer memory for recalling the initial labels, or thought that they might have been exposed to different initial labels. Thus, the participants were not aware of the existence of a minority in their group, but did notice potential differences between members of the groups and, subsequently, experienced difficulties in communicating with them.

2.3. Stimuli

We used a set of eight colorful alien images, adapted from Raviv and Arnon (2018) and designed by Noam Siegelman. Each alien was assigned a monosyllabic, bisyllabic, or trisyllabic nonsense label (fixed across participants) composed of a subset of letters from the unbiased participant's inventory (see Illustration 1, Panel C), that is, *kesip, esip, nus, nusa, aike, puak, nekuki, anap*. The initial language thus consisted of eight labels, and included a total of four instances for each unbiased letter, and five instances for each biased letter. The labels were selected such that two out of the 32 pairs of labels could only be distinguished by the use of the biased letter (nus**a**/nus, and **k**esip/esip). This design introduced an ambiguity, pushing the group to innovate to ensure successful understanding.

2.4. Procedure

The experiment consisted of five main parts (see Illustration 2). First, each participant was given two passive exposures to each of the eight alien images and their corresponding label, displayed together on the screen for 7 s.[3] Next, the participants were tested on their knowledge of this initial form-meaning mapping by being presented with each image once and asked to type the corresponding label. Then, the communication game began (see Illustration 3), and the participants played multiple rounds of repeated interactions in alternating pairs, switching between different communication partners in a systematic sequence (participant 1 is first paired with participant 2 in Round 1, followed by participant 3 in Round 2, and participant 4 in Round 3, continuing this pattern until Round 9) and taking turns as the producer or the guesser. Participants were informed that the unique objective of this

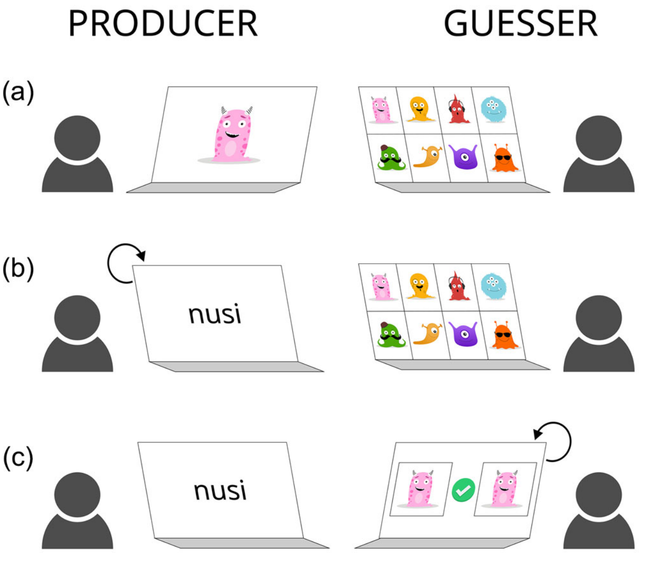

Illustration 3. Explanation of the communication game experiment. Panel A shows the producer being presented with an image and asked to come up with a label for it. In Panel B, the producer types the label and rotates their computer to present the label to the guesser. The guesser must then select which image they think is the correct one out of the eight possible images. In Panel C, after the guesser has chosen an image, feedback is presented showing the chosen image on the left, the image presented to the producer on the right, and an indicator in the middle displaying success or failure.

communication phase was to maximize their communicative success, and that they were not obliged to use the initial labels. This was done in order to mimic real-world communication, and to ensure participants are aware that they are allowed to deviate from the learned forms.

In a given interaction, the producer was presented with an image and had to type the label they believed corresponds to it. They then showed the written label to their partner by rotating their screens, and their partners had to guess which image the producer was referring to from the set of eight possible options (the correct image plus the seven other alternatives). After the guesser selected an image, both participants received feedback which included an indication of success or failure (i.e., whether the image was guessed correctly), as well as the correct versus selected image. Following the feedback, participants switched roles for the next interaction (the guesser became the producer and vice versa). In a given round, each paired participant acted as a producer and a guesser for all eight images, resulting in a total of 16 interactions per round. At the end of the round, participants switched partners and interacted with another member of the group. Each participant communicated with all others. The communication process consisted of nine rounds, allowing each participant to interact with each other member of the group three times. We compared two types of networks: *control* networks (where no participant is biased) to *heterogeneous* networks (with one biased participant) (see Illustration 1, Panel A). The order of stimuli presentation was randomized per round and per group. After nine communication rounds, participants were tested again on their knowledge of the language using a production completion test similar to the initial test they completed prior to interaction.

Finally, participants completed three additional tasks (see below) as well as a questionnaire about their subjective experience, and were debriefed by the experimenter. On average, the experiment lasted around 90 min.

2.5. Additional tasks

The study included three additional tasks. The first task is the Prosociality scale (Caprara et al., 2005), which we translated into Dutch. This questionnaire consists of 16 items that measure different aspects of prosocial behavior such as empathic concern, altruism, and volunteering using a 5-point Likert scale ranging from 1 (never) to 5 (very often). Participants' answers are then averaged to yield a total prosociality score between 1 and 5 (5 indicating high prosociality). The scale has demonstrated good reliability and validity in various studies (Caprara, Alessandri, & Eisenberg, 2012; Carlo & Randall, 2002; Luengo Luengo Kanacri et al., 2021). The full Prosociality questionnaire is provided in the Supplementary Materials.

For the second task, we adapted a version of the Dictator game as another way to measure prosociality, seeing as it provides a different (and maybe more accurate) measure of prosociality than self-report questionnaires (Böckler, Tusche, & Singer, 2016). Participants were presented with the following fictional situation (translated here from the Dutch original):

> Imagine that I give you an additional amount of 100 euros due to the excellent performance of your group during this experiment. Now, you have the choice to keep the full amount for yourself or to share it with the other participants. Since the other participants are not aware of this extra reward, the choice is entirely up to you. How much do you decide to share with the other participants?

Participants had the choice between (1) keeping the entire amount for themselves, (2) sharing the amount but keeping more for themselves, (3) equally splitting the amount between all group members, (4) sharing the amount but keeping less for themselves, or (5) giving away the entire amount. Thus, the final score ranges from 1 (not prosocial) to 5 (highly prosocial).

Third, participants had to perform a task-switching experiment adapted from Roger and Monsell's (1995) paradigm. This challenging task required participants to remember the rules of two different tasks and frequently switch between them: press Q or P for consonants versus vowels, or press Q or P for odd versus even numbers. The goal of this task was to quantify cognitive flexibility, which represents the capability to adapt to new situations and potentially serve as an indicator for predicting linguistic alignment. Additionally, this task also evaluated participants' working memory, a variable we measured to ensure the similarity of individuals within our control and homogeneous groups, while also exploring its potential influence on linguistic alignment. The task is presented as follows: a pair of a letter and a number (e.g., G6) is presented in a quadrant of the screen (either top-left, top-right, bottom-left, bottom-right; see Illustration 4). When the set of letter/number is presented in one of the top quadrants, the participant needs to respond solely based on the letter. However, when presented at the

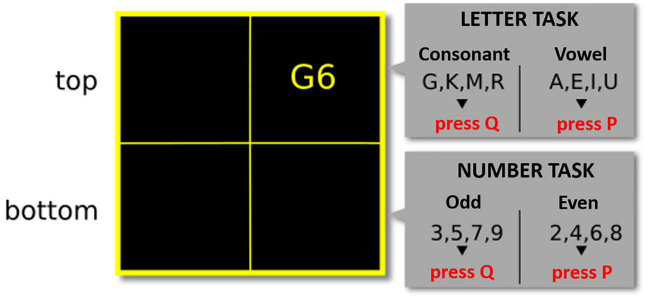

Illustration 4. Visual explanation of the task-switching experiment.

bottom, the participant needs to respond solely based on the number. The position of the set of letters/numbers alternates between the different positions of the quadrants, shifting sequentially from top-left to top-right, then to bottom-right, and finally to bottom-left. We extracted two measures from this task: the overall inverse efficiency (reaction time divided by accuracy), which measures the global working memory, and the task-switching cost, which indicated the difficulty of switching between tasks and was normalized by the total time (see Supplementary Materials for more information).

2.6. *Specific research questions*

In this research, we aim to answer the following research questions: (1) How did language evolve in the heterogeneous groups? (2) Has the language of the group adapted to the specificity of the biased participant? (3) Who adapts and when? Specifically, we seek to understand if the participants adapt to their partners in pair interactions, and if yes, what factors might predict the extent of the alignment.

2.7. *Measures*

Before turning to the measures aiming at answering specific research questions, it is essential to assess the participants' memory of the initial labels to make sure that there are no significant differences between our control and heterogeneous groups. To do so, we measured the ***learning accuracy*** and the ***initial production similarity*** after training. Learning accuracy was a binary measure, where a value of 1 indicates correct recall by the participant, while a value of 0 indicates the presence of at least one error. Given the biased participants' inability to produce certain labels, their responses were considered "correct" (1) if they replaced the biased letter with another letter or if they omitted it. The ***initial production similarity*** is the normalized Levenshtein distance between the initial labels and the participants' first production, reflecting a more fine-grained accuracy score that takes into account the proportion of incorrect letters in each production. By design, a participant who initially perfectly remembers the given labels will reach a perfect similarity (with a value of 1.0).

(1) We used a set of three different measures to address our first research question. First, ***communicative success*** indicates the success (1) or failure (0) of the inter-

action between two participants (as shown in Illustration 3, Panel C), aggregated by pairs. Second, ***convergence*** measures the similarity of the words generated for the same image in each round. It is computed using the pairwise normalized Levenshtein distance between the four labels produced for the same image at each round. Low convergence indicates that different words have been used to describe the same image, while high convergence suggests that participants used a similar label. Lastly, we examined the ***production similarity*** between participants' labels and the set of initial labels over time, by computing the average Levenshtein distance between participants' productions and the initial labels at a given time. By design, the biased participants cannot reach a perfect level of similarity because they lack two out of the eight letters present in the initial labels. For this reason, we present the results of production similarity in two ways for the heterogeneous groups: excluding versus including the biased participant.

(2) To answer the second research question, we refer to the ***frequency of the biased*** letters. For each round, we determine the total frequency of the biased letters ("k" and "a") out of all the letters used in that round, which we multiplied by 100 to obtain percentages. We also computed the frequency distribution of biased letters in the initial set of labels (*kesip, esip, nus, nusa, aike, puak, nekuki, anap*). We observe the frequency of biased letters for all participants in control groups, but we exclude the productions from the biased participant in heterogeneous groups.

(3) To identify the degree of accommodation across participants, we created several ***adaptability*** measures (full details in the Supplementary Materials). For the sake of simplicity, we focus here on one adaptability score. For each round, participant, and shape, we recorded participants' most recent label production, the preceding label produced by their partner, and the label they previously produced themselves. Subsequently, we calculated the normalized Levenshtein distance between the latest label and the partner's label (d_1), as well as the distance between the latest label and the participant's own previous label (d_2). Our adaptability measure, defined as $Adapt = 2 * (1 - d_1) + d_2$, predominantly reflects whether a participant employs the same label as their partner's just before (i.e., d_1). However, this value gains even more significance if the label significantly differs from their own previously produced label (d_2). We aggregated these values of adaptability per participant.

2.8. Analysis

All analyses were conducted using mixed-effect models implemented through the lmerTest package (Kuznetsova, Brockhoff, & Christensen, 2017), a package based on lme4 package (Bates, Mächler, Bolker, & Walker, 2015) in R (R Core Team, 2023). Interactions between fixed effects are considered in all models.

To answer the first research question *(1)*, the data were aggregated by pair (for *communicative success;* note that we also implemented a model based on individual trials, without aggregation by pair, which yielded similar but stronger results; see Supplementary Materials), by group (for *convergence*), and by participant (for *production similarity*). We built models with fixed factors for *round number* (coded as continuous, in reverse order so that Round 10 becomes the reference level[4]) and *group type* (control/heterogeneous, dummy coded, with control as the reference level). **Model 1** predicted *communicative success*, **model 2** predicted *convergence*, and **model 3** predicted *production similarity*. We controlled for the *group ID* as a random factor in all models. To specifically examine how the presence of a biased participant affected the unbiased participants, we also built the same models with a modification, splitting the productions of the participants in the heterogeneous group into two categories: the productions involving the biased participants and the productions involving only the unbiased participants. We compare three conditions: control, heterogeneous *with* the biased participant, and heterogeneous *without* the biased participant (dummy coded, with control as the reference level). These variations of the models will be referred to as **model 1 bis**, **model 2 bis**, and **model 3 bis**.

To answer (2), we investigated whether the language of the heterogeneous groups featured less biased letters after the participants communicated. To do so, we examined the interaction of *group type* (control/heterogeneous, dummy coded, with control as the reference level) with *round number* **(model 4)**. Here too, *round number* was coded as continuous, in reverse order so that Round 10 becomes the reference level. The random effects include *participants* nested within *groups* (note that attempting to include the round as a random slope caused the model to fail to converge). Additional models, which either aggregated data by groups or focused solely on the *testing moments* (beginning vs. end of the experiment), support the findings of model 4, and can be found in the Supplementary Materials.

Regarding question *(3),* to predict the frequency of biased letters at the pair level, we aggregated the frequency of biased letters in two types of situations: when unbiased participants interact with the biased participant, and when they interact with another unbiased participant (this variable is called "Communicate With Biased"). **Model 5** predicted the *mean frequency of biased letters* in heterogeneous groups using "*Communicate With Biased*" ("yes"/"no," dummy coded, with "no" as the reference level) as fixed factors. In this model, we also have the *group ID* and the *participant's unique ID* as random factors, since "Communicate With Biased" is a within-participants variable. Then, in order to understand which factors affect adaptability, we predicted adaptability using *age*, *gender*, *working memory*, *cognitive flexibility*, *prosociality*, and learning success measures (*accuracy learning* and *initial production similarity*) as fixed effects (**model 6**) and *group ID* as a random effect: all variables are centered except for gender (dummy coded, with female as the reference level). Notably, this is the only model in which we consider only the main effects of the fixed factors without their interactions. We opted for this approach because including interactions resulted in a model that was overly complex, and we lacked any prior hypotheses regarding the effects of these interactions.

3. Results

We tested seven heterogeneous groups and seven control groups. The distribution of age, gender, and the other measures is similar between the two types of groups. On average,

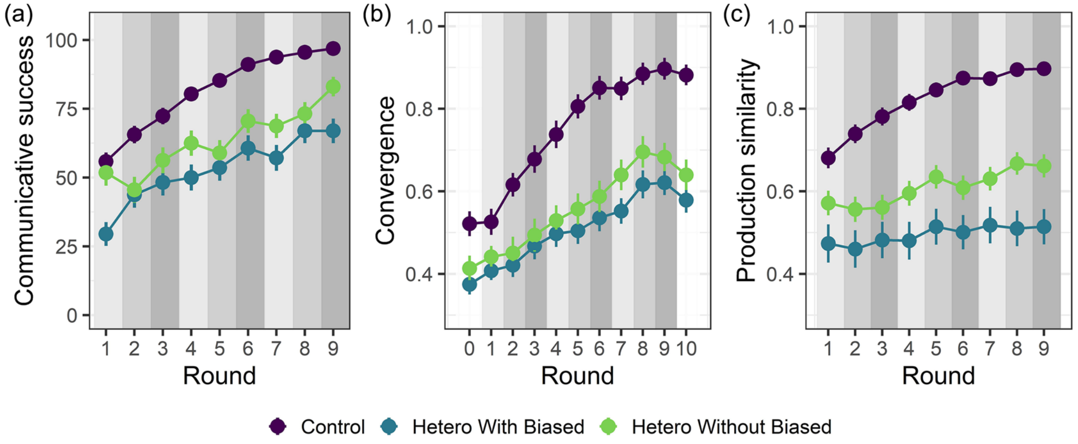

Fig. 1. The presence of a biased participant impacts communication in the heterogeneous groups: Panel A. Evolution of communicative success (% of successful interactions) from Round 1 to 9. The dark purple dots and error bars represent, respectively, the average and standard-error communicative success for all pairs in all control groups. In heterogeneous groups, we represent the average communicative success for pairs that did not include the biased participant (in green) or for pairs that included the biased participant (in blue). Panel B. Evolution of mean convergence. The y-axis shows the pairwise normalized Levenshtein distance between all words within the control group (in dark purple), heterogeneous group with productions from all participants (in blue), and heterogeneous groups excluding productions from the biased participant (in green). This includes convergence for the testing productions before (Round = 0) and after (Round = 10) communication game. Panel C. Evolution of the production similarity with the initial labels with time, for control groups (in dark purple), as well as for unbiased (in green) and biased (in blue) participants in heterogeneous groups.

participants remembered half of the initial labels (approximately 47% of correct recall, with a standard-deviation of 27%, indicating strong differences in participants' ability to recall the labels after the first exposure). Biased participants, when they accurately recall the initial labels, either substituted the biased letters with another unbiased letter (48%) or omitted the biased letters altogether (52%).

3.1. *How did language evolve in the heterogeneous groups?*

Communicative success was lower in heterogeneous groups compared to control groups (model 1: $\beta_{GroupTypeHetero} = -27.9 \pm 7.3$, where 7.3 is the standard error, $p = .0015$, see Fig. 1A) and increased with time for both types of groups (model 1: $\beta_{Round} = -5.1 \pm 0.5$, $p < 2{\cdot}10^{-16}$; Fig. 1A). Similarly, *convergence* increased with time for both groups (model 2: $\beta_{Round} = -0.04 \pm 0.002$, $p < 2{\cdot}10^{-16}$; Fig. 1B) and was lower for heterogeneous groups (model 2: $\beta_{GroupTypeHetero} = -0.33 \pm 0.06$, $p = 6.01{\cdot}10^{-5}$; Fig. 1B). Furthermore, *convergence* increased faster for control groups compared to heterogeneous groups (model 2: $\beta_{GroupTypeHet:Round} = 0.02 \pm 0.004$, $p = 3.0{\cdot}10^{-6}$; Fig. 1B). The greater convergence observed in the control groups could be due to their tendency to align with the initial labels, even if these labels were not initially recalled (refer to Fig. 1C). In contrast, the heterogeneous groups consistently employed labels that deviated from the initial ones (significant interaction between

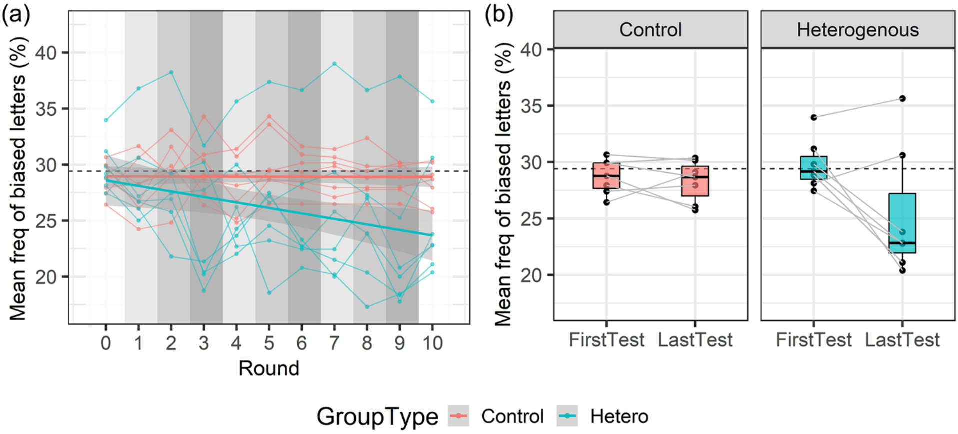

Fig. 2. Most heterogeneous groups progressively avoid the biased letters: In both panels, we observe the evolution of the mean frequency of biased letters (in %, see Method for more details on this measure's calculation) over time. Control groups are represented in red, while heterogeneous groups are shown in blue. The black dashed line represents the frequency of biased letters in the initial labels. Panel A: Aggregated data by group (with each line representing one group) for rounds 1−9. The thick bold line represents the linear regression passing through these points. Panel B Left: change in the mean frequency of biased letters before communication (first test) and after communication (last test) for the control groups. Panel B Right: the same for the heterogeneous groups.

the type of group and round in model 3: $\beta_{GroupTypeHetero:Round} = -0.01 \pm 0.004$, $p = 6.4 \cdot 10^{-5}$; Fig. 1C).[5]

One can imagine that the lower *communicative success* and *convergence* within heterogeneous groups were solely due to the outputs produced by the biased participant. However, it seems that their impact went beyond that, and that the biased participants influenced the overall performance of the entire group. In models distinguishing between productions containing or excluding the biased participant in heterogeneous groups, even pairs including only unbiased participants also had lower *communicative success* compared to control groups (model 1 bis: $\beta_{HeteroWithoutBiased} = -23 \pm 7.7$, $p = .0074$), *convergence* (model 2 bis: $\beta_{HeteroWithoutBiased} = -0.25 \pm 0.06$, $p = .0009$), and *production similarity* (model 3 bis: $\beta_{HeteroWithoutBiased} = -0.27 \pm 0.06$, $p = .0005$) compared to control groups. This is also visible in Fig. 1.

3.2. *Has the language of the group adapted to the specificity of the biased participant?*

We analyzed the frequency of biased letters used across the rounds, including the first and the final testing phase. The frequency of the biased letters in the initial labels is of 29.4% (10 letters out of 34, see Illustration 1C and see dotted line in Figs. 2 and 3). Dropping one biased letter from the set of initial labels translates into a decrease of about 2% of this frequency (9 out of 33, 27.3%). Control groups exhibited the expected pattern, with fluctuations around an average constant frequency of biased letters (Fig. 2, in red) throughout the experiment. Among the heterogeneous groups, five out of the seven groups adjusted their language in

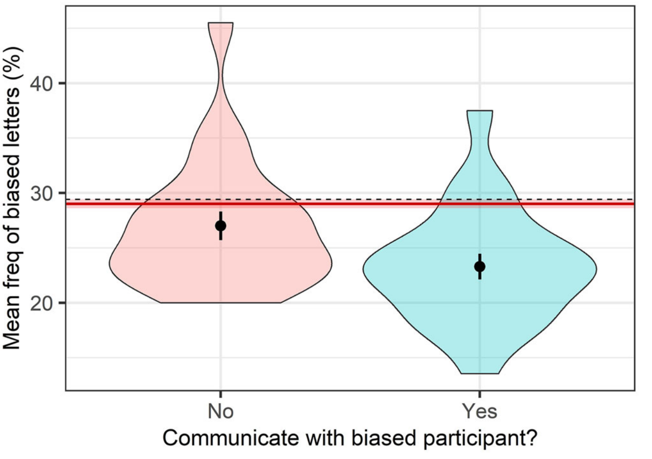

Fig. 3. Participants specifically adapt to the biased participant: Mean and standard error frequency of biased letters within heterogeneous groups, when unbiased participants are interacting with other unbiased participants (left) or with the biased participant (right). The black dashed line represents the biased letter frequency in the initial labels. The red bold line corresponds to the average frequency of the biased letters in control groups, accompanied by standard error ribbons.

response to the presence of biased participants (Fig. 2, Panel B, in blue). Notably, the final language of these groups contained fewer biased letters than their initial language. Thus, individuals in heterogeneous groups (excluding the productions from the biased participant) use on average fewer biased letters than individuals in control groups (model 4: $\beta_{GroupTypeHet} = -5.2 \pm 1.95$, $p = .0192$), and their frequency of biased letters decreases more with time than for individuals in the control groups (model 4: $\beta_{Round*GroupTypeHet} = 0.47 \pm 0.10, p = 6.93{\cdot}10^{-6}$).

These adaptations took various forms. Unbiased participants sometimes adopted the label introduced by the biased participant, who either substituted the biased letter, removed it, or used a new label (in 10% of the occurrences where the unbiased participants did adapt). Most of the time, unbiased participants chose an alternative label that was distinct from the one used by the biased participant, yet without including any biased letters (48%) or with fewer biased letters than in the initial labels (going from two to one biased letter; 43%). When they did not adapt, they either used the initial labels (64% of the occurrences where the unbiased participants did *not* adapt) or chose an alternative label that includes as many (or more) biased letters as the initial labels (36%).

Why did two of these groups not adapt? Notably, one of these two groups exhibited a distinct pattern: three participants, including the biased individual, remembered very well the initial labels in the very first test (notably, it was the only heterogeneous group where three participants had more than 50% of accuracy, including two participants who had more than 80% accurate recall). In this group, the biased participant consistently replaced the biased letter with other unbiased letters, and the group quickly converged on the initial labels with slight and consistent variations on the biased participant's side. This stable synonymy did not hamper their communicative success. The reasons behind the lack of adaptation in the other group are less clear. However, it is interesting to note that two out of the four participants in this group also demonstrated very high recall for the initial labels (around 75%). By contrast, in other heterogeneous groups, it was rare to find more than one participant achieving over 50% accuracy. This situation raises intriguing questions: What factors determine who adapts and who does not? What personal attributes of group members contribute to the adaptation of the group's language?

3.3. *Who adapts and when?*

A clear pattern of adaptation emerges at the pair level: when unbiased participants communicated with the biased participant, they noticeably decreased their use of the biased letters (model 5: $\beta_{ComBiased} = -1.86 \pm 0.37$, $p = 7.54\text{e-}05$; see Fig. 3).

We also aim to understand which participants were most likely to adapt and align with their partner's label using our adaptability score. This score is not specifically related to the presence or absence of biased letters, but is more generally associated with a participant's inclination to adopt their partner's productions, particularly when it contradicts their a-priori preference. Our results suggest that factors such as gender, age, and various metrics derived from supplementary tasks (including working memory, cognitive flexibility, and prosociality) did not predict the participants' likelihood of adaptation in our sample. Instead, the propensity of participants to align with their partners seems to be influenced solely by the participant's initial representation of the language, assessed by the production similarity (model 6: $\beta_{ProdSimilarity} = 0.51 \pm 0.24$, $p = .043$) and, to a lesser extent, the accuracy of learning (model 6: $\beta_{AccLearning} = -0.38 \pm 0.19$, $p = .050$). Thus, the stronger the participants' recall of the initial labels, the less inclined they were to align. We also found that participants who have higher scores of adaptability also tend to have higher communicative success on average (Spearman correlation between the two measures aggregated for each participant: $\rho = 0.50$, $S = 15$, $p = 7.7{\cdot}10^{-5}$).

3.4. *Summary of results*

To summarize our results, we found that languages evolved differently in groups where a minority had a bias affecting their language. First, heterogeneous groups showed lower communicative success and overall less and slower convergence toward shared labels—a result that can be explained by more deviation from the initial labels in such groups. In addition, we observed strong partner-specific alignment, and a partial carry-over to the group-level: participants typically adapted to the peculiarities of the biased individual during pairwise interaction, and five out of the seven groups showed a group-level adaptation of their language. They did so by reducing the use of the biased letters in all interactions, even in those that did not include the biased participant. Notably, the degree of this population-level adaptation was negatively correlated with the initial individuals' learning accuracy. In the two groups that did not adapt their languages, most participants showed a very high recollection of the initial labels learnt during training. Moreover, on an individual level, learning accuracy of the initial labels significantly predicted participants' tendency to align with others, rather than their age, gender, prosociality score, or cognitive flexibility. Together, these results suggest that individuals are able to adapt their languages to fit the variations introduced by a minority and that such pairwise variations *can* percolate to the rest of the community, and also accumulate over time—causing languages to change from their initial state in the presence of a minority.

4. Discussion

This study investigated how language evolves in heterogeneous versus homogeneous groups, and whether a production bias present in the minority could influence the language of the whole group. Overall, we observed differences in the languages used by the control (homogeneous) and the heterogeneous groups within this experimental setting. Our results suggest that heterogeneous groups exhibit a lower convergence on shared labels, notably characterized by the use of novel and/or synonymous labels for the same image. In contrast, control groups rapidly align on the initial labels they learned, despite not accurately recalling them on average. This is likely because, despite imperfect recall, participants are still able to recognize the initial labels when their partner produces them, leading to successful communication and reactivation of the memory for those labels. This raises the following question: does the higher linguistic variability in heterogeneous groups stem solely from the fact it takes more time for them to converge onto shared labels, or is it an inherent trait of heterogeneous populations that persists over time? Put differently, if we were to allow both heterogeneous and homogeneous groups to interact for 100 rounds, would the language of heterogeneous groups still exhibit greater variation than the control groups after communicating? Importantly, these two scenarios are not mutually exclusive. Fig. 1B's trajectory seems to indicate that language has not yet stabilized even after nine rounds of communication, implying that the convergence might require more interactions in heterogeneous populations. Previous computational work (Josserand et al., 2021; Josserand et al., 2024) supports both proposals, as it shows that language takes more time to stabilize in heterogeneous compared to homogeneous populations, and that the presence of individual differences contributes to higher language variability. This notion finds further support in sociolinguistic research, which underscores that linguistic diversity is high within groups encompassing a mix of social classes or diverse ethnic backgrounds (Bright, 2017; Labov, 1986).[6] Naturally, individual differences are omnipresent, both in real-world scenarios and within our experiment: our participants showcased a wide array of differences, spanning from personality traits, attitude toward language, to various other unmeasured cognitive abilities that could potentially impact language. Nevertheless, participants in our control group had access to the same letters on the keyboard and were homogeneous in that respect.

Another intriguing facet of these findings lies in the similar pattern exhibited by biased and unbiased participants in heterogeneous groups. The influence of the bias in the heterogeneous groups spreads beyond interactions involving the biased participants, as the interactions between the unbiased participants also exhibit lower convergence and communicative success rates. This suggests that the introduction of a biased minority exerts a broader impact on the overall convergence of the language of the whole group, including the unbiased majority. Thus, the presence of individual differences within populations may act as a catalyst for the emergence of innovative linguistic elements, thereby contributing to the broader landscape of cross-linguistic diversity. This aligns with applied research indicating a positive correlation between long-term innovation and diversity in various aspects such as gender (Nielsen, Bloch, & Schiebinger, 2018), age (only if the age groups are heterogeneous and not polarized; Mothe & Nguyen-Thi, 2021), and cultural diversity (Scharf, Foster, & Behdinan, 2014).

The languages generated by the heterogeneous groups also changed in response to the idiosyncrasies of the biased participants (though not simply via adopting their productions), as they contained significantly fewer biased letters after communication than initially—reflecting the lingering effect of the minority on the majority language. The changes eventually percolated to the rest of the community and affected the language of the whole group. The underlying mechanisms remain uncertain, but they may be associated with the variability of input received or memory interference caused by biased participants, resulting in the forgetting of initial labels. The fact that the biased letters were important to disambiguate relevant distinctions in the initial language plays a crucial role in explaining the strategies used by participants to overcome their deletion when interacting with the biased participant, and may help unpack why and how certain changes percolated in the group. While this phenomenon was subtle, it aligns with theories of cultural evolution—which propose that languages gradually adjust to the biases of its speakers due to their pervasive daily usage and transmission across generations, but that these adjustments likely span multiple generations and take time to accumulate (Dediu et al., 2017; Dediu & Moisik, 2019; Kirby et al., 2007). We suggest that with more time to interact and/or larger lexicons that are more demanding on participants' memory, the effect would become even stronger.

Notably, two out of the seven groups exhibit a final language that has *not* adapted to the biased participant. This absence of systematic adaptation is not surprising given notable real-world examples: while emergent sign languages demonstrate that majority language users can adapt to the biases of the minority, conversely, many real-world languages have not adapted to the idiosyncrasies of a minority. For instance, Romance languages still often feature the alveolar trill sound [r] in their sound systems even when a small percentage of the population is unable to produce it (Anselme, 2022). While these differences in adaptation may stem from a complex interplay of factors (Bright, 2017), this dynamic may also be random. Notably, the unique characteristics of individuals within a group might also add to the complex nature of this process. For example, people might adapt more to a prestigious minority than to a stigmatized one (Bloomfield, 1933; Fagyal et al., 2010). Moreover, the proportion of biased individuals within each could also have significant implications: for example, Would we see a weaker or greater tendency to accommodate in groups with smaller or larger minorities? And would a lower ratio result in more nonadaptive groups, or simply a reduced degree of adaptation? And is there a cutoff point beyond which individual differences no longer affect the group's language? These questions could be explored in future work, testing different-sized groups with different proportions of biased individuals (e.g., 1:8). Computational models suggest that lower ratios of biased agents lead to weaker adaptations (Josserand et al., 2021).

To understand the individual dynamics of adaptation, we zoomed in to observe what happens during interaction at the pair level. First, it is important to keep in mind that unbiased participants in the heterogeneous groups were unaware of the presence of a production bias in one of the other participants within the group. This lack of awareness thus did not predispose them toward overtly prosocial or rejecting behaviors. Nevertheless, participants in our laboratory setup demonstrated a strong propensity to adopt similar linguistic variants as their conversation partners, both biased and unbiased, which supports previous observations of linguistic alignment in real-world interaction (Danescu-Niculescu-Mizil, Gamon, & Dumais, 2011; Niederhoffer & Pennebaker, 2002). Participants who aligned more with their partners had greater communicative success, supporting the interactive alignment framework, which suggests that conversation is facilitated by a process that aligns linguistic representations between partners (Garrod & Pickering, 2004). However, while interacting with a biased participant did lead to changes in the unbiased participants' languages (with a global reduction in biased letter frequency), it was only in a minority of cases that the exact labels produced by the biased participant were taken over by the unbiased participants. Thus, the current findings suggest a more complex adaptation process than simply adopting the labels produced by the biased participant, and further research is needed to explore the precise mechanisms through which individuals adapt to the idiosyncrasies of a minority.

Interestingly, unlike some previous studies where the propensity to align was influenced by factors such as gender, age, or prosociality, our findings did not find such associations. Several reasons may contribute to this: First, our groups were relatively homogenous in terms of age and gender demographics: all participants were between 18 and 37 years old, and 68% of them were female. Similarly, our participants showed very similar (and relatively high) prosociality scores, with 95% of participants having prosociality scores between 3 and 4.5 on the 0−5 prosociality scale. Thus, our sample might not have been sufficiently variable on these dimensions to allow such associations. It is important to note that, since our goal was to examine implicit processes of language adaptation and change, our participants were not made aware of the bias of one of the members in their group. However, this is not always the case in real-life scenarios: people often consciously devise strategies to accommodate individuals with observable linguistic deficiencies, poor fluency, and speech idiosyncrasies, for example, as is the case with child- or foreigner-directed speech (Berdicevskis, 2020; Ferguson & Peterson, 2002; Schick et al., 2022; Simmons-Mackie, 2018). However, many idiosyncrasies that are perceived as pathological or socially "unacceptable" can carry a negative societal value that might consciously deter people from adopting them. Future research could explore how increasing participants' awareness of a biased individual affects these adaptation processes.

However, a notable predictor of the participants' tendency to align with their partners' linguistic productions was their ability to recall the initial labels: Participants who demonstrated a strong memory recall of the initial labels showed less adaptation to their partner's language. A similar pattern was found for one of the two groups that did not show signs of adaptation, in which participants showed an exceptionally good memory of the initial labels. The link between better learning and reduced alignment could be potentially attributed to the participants' perception of the initial labels as representing the "ground truth" of the language. This idea aligns closely with findings from computational models, which suggest that the stronger a person's belief is in an initial language, the less likely they are to deviate from it and introduce variation (Josserand et al., 2024). These findings could be potentially related to the enforcing of linguistic policies, for example, in situations where a language is perceived to have a "correct" or "acceptable" form, it may be likely to observe a reduced inclination for adopting new linguistic innovations. For example, the institution *Académie française* in France established strict linguistic norms, especially oriented toward the orthographic system, which may prevent people from adopting orthographic innovations (Walter, 2016). This phenomenon also extends to spoken language, where explicit language ideologies emerge as people propose statements regarding their own language skills, behaviors, or those of others (Pakendorf, Dobrushina, & Khanina, 2021).

Together, this study provides an important step in understanding the weight of individual constraints on the formation of collective language characteristics, and emphasizes the importance of accounting for individual differences within a group to better understand language evolution. A compelling avenue for future investigation lies in studying potential interactions between group size and individual differences, seeing as the impact of a single individual may be bigger in a smaller community, but washed away in a larger community. Indeed, previous work shows that smaller groups are more sensitive to random idiosyncrasies such as those introduced by a single participant, while larger groups seem to be more immune to such drift (Raviv et al., 2019b). Similarly, exploring diverse network structures where the biased participant occupies either a central or a peripheral role could shed light on how the spread of biased variants is influenced by network topology (Raviv et al., 2020). For example, having the biased participant in a central position in the network (representing an influential figure, for example), can potentially lead to them having a stronger impact on the entire community. Moreover, extending the exploration of biases in language production to encompass biases in perception could yield valuable insights, as research suggests that individual differences in color perception may affect language evolution (Josserand, Meeussen, Majid, & Dediu, 2021; Lindsey & Brown, 2002). Finally, since our findings highlight a notable discrepancy in the language used at the pair level and at the group level, they underscore the importance of employing hierarchical Bayesian models which can account for such differences in future modeling research (Hawkins et al., 2023; Thompson, Raviv, & Kirby, 2019). While the current study looked at language, we believe that the main conclusions, namely, that group-level patterns are shaped by individual differences, can be applied across other diverse fields of cultural evolution, including the evolution of tool use and cultural norms.

4.1. Conclusions

Our findings underscore that introducing a minority with different language abilities within a group influences the trajectory of language evolution. Notably, the presence of individual variations fostered the emergence of novel linguistic elements. Furthermore, most individuals adapt their languages during paired interactions to accommodate variations introduced by a minority. In five out of the seven groups, these adaptations spread to the group level with a notable reduction in the use of the letters that one of the participants was unable to produce, effectively shifting the language of the community to support the minority. The extent of this adaptation is intertwined with how strongly participants remembered the initial labels, suggesting that a stronger attachment to a conventionalized form leads to lower adjustment to the idiosyncrasies of a biased minority. Together, our results provide compelling evidence that individual differences can diffuse throughout the broader community and accumulate over time, ultimately driving changes in languages. Hence, we believe it is crucial to integrate individual differences in future work and acknowledge them as significant factors that may influence the course of language evolution.

Author contributions

Mathilde Josserand: Methodology, software, validation, formal analysis, investigation, writing—original draft, project administration. **François Pellegrino**: Methodology, writing—review and editing, supervision. **Oxana Grosseck**: Investigation, writing—review and editing, project administration. **Dan Dediu**: Conceptualization, methodology, writing—review and editing, supervision. **Limor Raviv**: Conceptualization, methodology, resources, writing—review and editing, supervision, project administration, funding acquisition.

Acknowledgments

We also thank Noam Siegelman for sharing the set of alien images, and Tomas Lehecka and Shiri Lev-Ari for their contributions during the initial discussions of this project.

Funding information

This work was supported by the Van Gogh program (PHC VAN GOGH; project number 49212TM).

Data availability statement

The program used for the experiment, the data, and the code used to generate the analysis and the Supplementary Materials are available at https://github.com/mathjoss/ExpCommunityVariation.

Notes

1 This is also known under *linguistic style adaptation* or *linguistic style matching* concepts, even if these refer to slightly different concepts (Postmus, 2023).
2 Only one participant was a German native speaker but has a C2 level in Dutch and speaks it on a daily basis for more than 10 years.
3 This design was selected following an extensive pilot phase in order to balance between initial learning, the duration of the experiment, and the potential for flexibility. Other designs (in particular, those with more or less initial training) can be considered in follow-up experiments.

4 We performed this transformation since we are interested in the main effect of group at the final round.
5 Please refer to the Supplementary Materials for the analysis of stability evolution across rounds.
6 Of course, we acknowledge that individual differences are not the sole determinant influencing emergent linguistic variation, as relatively homogenous populations sharing similar contexts can also exhibit substantial lexical diversity (Mudd, de Vos, & de Boer, 2022).

References

Allen, J. S., Miller, J. L., & DeSteno, D. (2003). Individual talker differences in voice-onset-time. *Journal of the Acoustical Society of America*, *113*(1), 544–552. https://doi.org/10.1121/1.1528172

Anselme, R. (2022). *La représentation du trill en typologie : de l'acoustique à l'inventaire phonémique* [PhD Thesis]. Université Lumière - Lyon II.

Bates, D., Mächler, M., Bolker, B., & Walker, S. (2015). Fitting linear mixed-effects models using lme4. *Journal of Statistical Software*, *67*(1), 1–48. https://doi.org/10.18637/jss.v067.i01

Berdicevskis, A. (2020). Foreigner-directed speech is simpler than native-directed: Evidence from social media. In D. Bamman, D. Hovy, D. Jurgens, B. O'Connor, & S. Volkova (Eds.), *Proceedings of the Fourth Workshop on Natural Language Processing and Computational Social Science* (pp. 163–172). Association for Computational Linguistics. https://doi.org/10.18653/v1/2020.nlpcss-1.18

Bilous, F. R., & Krauss, R. M. (1988). Dominance and accommodation in the conversational behaviours of same- and mixed-gender dyads. *Language & Communication*, *8*(3), 183–194. https://doi.org/10.1016/0271-5309(88)90016-X

Blasi, D. E., Moran, S., Moisik, S. R., Widmer, P., Dediu, D., & Bickel, B. (2019). Human sound systems are shaped by post-Neolithic changes in bite configuration. *Science*, *363*(6432), eaav3218. https://doi.org/10.1126/science.aav3218

Bloomfield, L. (1933). *Language*. New York: Holt.

Böckler, A., Tusche, A., & Singer, T. (2016). The structure of human prosociality: Differentiating altruistically motivated, norm motivated, strategically motivated, and self-reported prosocial behavior. *Social Psychological and Personality Science*, *7*(6), 530–541. https://doi.org/10.1177/1948550616639650

Bortfeld, H., & Brennan, S. E. (1997). Use and acquisition of idiomatic expressions in referring by native and non-native speakers. *Discourse Processes*, *23*(2), 119–147. https://doi.org/10.1080/01638537709544986

Branigan, H. P., Pickering, M. J., Pearson, J., & McLean, J. F. (2010). Linguistic alignment between people and computers. *Journal of Pragmatics*, *42*(9), 2355–2368. https://doi.org/10.1016/j.pragma.2009.12.012

Brennan, S. E., & Clark, H. H. (1996). Conceptual pacts and lexical choice in conversation. *Journal of Experimental Psychology: Learning, Memory, and Cognition*, *22*(6), 1482–1493. https://doi.org/10.1037/0278-7393.22.6.1482

Bright, W. (2017). Social factors in language change. In F. Coulmas (Ed.), *The handbook of sociolinguistics* (pp. 81–91). Blackwell Publishing Ltd. https://doi.org/10.1002/9781405166256.ch5

Byrd, D. (1992). Preliminary results on speaker-dependent variation in the TIMIT database. *Journal of the Acoustical Society of America*, *92*(1), 593–596. https://doi.org/10.1121/1.404271

Caprara, G. V., Alessandri, G., & Eisenberg, N. (2012). Prosociality: The contribution of traits, values, and self-efficacy beliefs. *Journal of Personality and Social Psychology*, *102*(6), 1289–1303. https://doi.org/10.1037/a0025626

Caprara, G. V., Steca, P., Zelli, A., & Capanna, C. (2005). A new scale for measuring adults' prosocialness. *European Journal of Psychological Assessment*, *21*(2), 77–89.

Carlo, G., & Randall, B. A. (2002). The development of a measure of prosocial behaviors for late adolescents. *Journal of Youth and Adolescence*, *31*(1), 31–44. https://doi.org/10.1023/A:1014033032440

Carr, J. W., Smith, K., Culbertson, J., & Kirby, S. (2020). Simplicity and informativeness in semantic category systems. *Cognition*, *202*, 104289. https://doi.org/10.1016/j.cognition.2020.104289

Chartrand, T. L., & Bargh, J. A. (1999). The chameleon effect: The perception–behavior link and social interaction. *Journal of Personality and Social Psychology*, *76*(6), 893–910. https://doi.org/10.1037/0022-3514.76.6.893

Clark, H. H. (1996). *Using language*. Cambridge University Press. https://doi.org/10.1017/CBO9780511620539

Cools, R. (2015). Neuropsychopharmacology of cognitive flexibility. In A. W. Toga (Ed.), *Brain mapping* (pp. 349–353). Academic Press. https://doi.org/10.1016/B978-0-12-397025-1.00253-0

Culbertson, J. (2012). Typological universals as reflections of biased learning: Evidence from artificial language learning. *Language and Linguistics Compass*, *6*(5), 310–329. https://doi.org/10.1002/lnc3.338

Danescu-Niculescu-Mizil, C., Gamon, M., & Dumais, S. (2011). Mark my words!: Linguistic style accommodation in social media. In *Proceedings of the 20th International Conference on World Wide Web - WWW '11* (p. 745). https://doi.org/10.1145/1963405.1963509

De Busser, R. (2015). *The influence of social, cultural, and natural factors on language structure: An overview*. https://doi.org/10.1075/clscc.6.01bus

De Vos, C. (2011). A signers' village in Bali, Indonesia. *Minpaku Anthropology Newsletter*, *33*, 4–5.

Dediu, D., Janssen, R., & Moisik, S. R. (2017). Language is not isolated from its wider environment: Vocal tract influences on the evolution of speech and language. *Language & Communication*, *54*, 9–20. https://doi.org/10.1016/j.langcom.2016.10.002

Dediu, D., Janssen, R., & Moisik, S. R. (2019). Weak biases emerging from vocal tract anatomy shape the repeated transmission of vowels. *Nature Human Behaviour*, *3*, 1107–1115. https://doi.org/10.1038/s41562-019-0663-x

Dediu, D., & Moisik, S. R. (2019). Biology matters: Variation in vocal tract anatomy and language. *Selected Papers on Theoretical and Applied Linguistics*, *23*, 19–33.

Everett, C., Blasi, D. E., & Roberts, S. G. (2015). Climate, vocal folds, and tonal languages: Connecting the physiological and geographic dots. *Proceedings of the National Academy of Sciences*, *112*(5), 1322–1337. https://doi.org/10.1073/pnas.1417413112

Fagyal, Z., Swarup, S., Escobar, A. M., Gasser, L., & Lakkaraju, K. (2010). Centers and peripheries: Network roles in language change. *Lingua*, *120*(8), 2061–2079. https://doi.org/10.1016/j.lingua.2010.02.001

Fehér, O., Ritt, N., & Smith, K. (2019). Asymmetric accommodation during interaction leads to the regularisation of linguistic variants. *Journal of Memory and Language*, *109*, 104036. https://doi.org/10.1016/j.jml.2019.104036

Fehér, O., Wonnacott, E., & Smith, K. (2016). Structural priming in artificial languages and the regularisation of unpredictable variation. *Journal of Memory and Language*, *91*, 158–180. https://doi.org/10.1016/j.jml.2016.06.002

Ferdinand, V., & Zuidema, W. (2009). Thomas' theorem meets Bayes' rule: A model of the iterated learning of language. In *Proceedings of the 31st Annual Conference of the Cognitive Science Society* (pp. 1786–1791). Retrieved from https://csjarchive.cogsci.rpi.edu/Proceedings/2009/papers/376/paper376.pdf

Ferguson, A., & Peterson, P. (2002). Intonation in partner accommodation for aphasia: A descriptive single case study. *Journal of Communication Disorders*, *35*(1), 11–30. https://doi.org/10.1016/s0021-9924(01)00071-5

Fillmore, C. J., Kempler, D., & Wang, W. S. (2014). *Individual differences in language ability and language behavior*. Academic Press.

Foulkes, P., & Docherty, G. (2006). The social life of phonetics and phonology. *Journal of Phonetics*, *34*(4), 409–438. https://doi.org/10.1016/j.wocn.2005.08.002

Fox, M. (2008). *Talking hands: What sign language reveals about the mind* (Reprint édition). Simon & Schuster.

Frank, S., & Smith, K. (2018). *A model of linguistic accommodation leading to language simplification* [Preprint]. PsyArXiv. https://doi.org/10.31234/osf.io/4ynwu

Garrod, S., & Pickering, M. J. (2004). Why is conversation so easy? *Trends in Cognitive Sciences*, *8*(1), 8–11. https://doi.org/10.1016/j.tics.2003.10.016

Giles, H., Coupland, N., & Coupland, J. (1991). Accommodation theory: Communication, context, and consequence. In H. Giles, J. Couplang, & N. Couplang (Eds.), *Contexts of accommodation: Developments in applied sociolinguistics* (pp. 1–68). chapter, Cambridge: Cambridge University Press. https://doi.org/10.1017/CBO9780511663673.001

Gnisci, A. (2005). Sequential strategies of accommodation: A new method in courtroom. *British Journal of Social Psychology*, *44*Pt (4), 621–643. https://doi.org/10.1348/014466604x16363

Griffiths, T. L., & Kalish, M. L. (2007). Language evolution by iterated learning with Bayesian agents. *Cognitive Science*, *31*(3), 441–480. https://doi.org/10.1080/15326900701326576

Hawkins, R. D., Franke, M., Frank, M. C., Goldberg, A. E., Smith, K., Griffiths, T. L., & Goodman, N. D. (2023). From partners to populations: A hierarchical Bayesian account of coordination and convention. *Psychological Review*, *130*(4), 977–1016. https://doi.org/10.1037/rev0000348

Jacques, S., & Zelazo, P. D. (2005). Language and the development of cognitive flexibility: Implications for theory of mind. In J. W. Astington & J. A. Baird (Eds.), *Why language matters for theory of mind* (pp. 144–162). Oxford University Press. https://doi.org/10.1093/acprof:oSo/9780195159912.003.0008

Johnson, K., Ladefoged, P., & Lindau, M. (1993). Individual differences in vowel production. *Journal of the Acoustical Society of America*, *94*2 Pt(1), 701–714. https://doi.org/10.1121/1.406887

Jones, E., Gallois, C., Callan, V., & Barker, M. (1999). Strategies of accommodation: Development of a coding system for conversational interaction. *Journal of Language and Social Psychology*, *18*(2), 123–151. https://doi.org/10.1177/0261927x99018002001

Jones, S., Cotterill, R., Dedney, N., Muir, K., & Joinson, A. (2014). Finding Zelig in text: A measure for normalising linguistic accommodation. In *COLING 2014*.

Josserand, M., Allassonnière-Tang, M., Pellegrino, F., & Dediu, D. (2021). Interindividual variation refuses to go away: A Bayesian computer model of language change in communicative networks. *Frontiers in Psychology*, *12*, 626118. https://doi.org/10.3389/fpsyg.2021.626118

Josserand, M., Allassonnière-Tang, M., Pellegrino, F., Dediu, D., & de Boer, B. (2024). How network structure shapes languages: Disentangling the factors driving variation in communicative agents. *Cognitive Science*, *48*(4), e13439. https://doi.org/10.1111/cogs.13439

Josserand, M., Meeussen, E., Majid, A., & Dediu, D. (2021). Environment and culture shape both the colour lexicon and the genetics of colour perception. *Scientific Reports*, *11*(1), 19095. https://doi.org/10.1038/s41598-021-98550-3

Kapnoula, E. E. (2016). *Individual differences in speech perception: Sources, functions, and consequences of phoneme categorization gradiency* [Doctor of Philosophy], University of Iowa. https://doi.org/10.17077/etd.feubmitk

Kirby, S., Cornish, H., & Smith, K. (2008). Cumulative cultural evolution in the laboratory: An experimental approach to the origins of structure in human language. *Proceedings of the National Academy of Sciences*, *105*(31), 10681–10686. https://doi.org/10.1073/pnas.0707835105

Kirby, S., Dowman, M., & Griffiths, T. L. (2007). Innateness and culture in the evolution of language. *Proceedings of the National Academy of Sciences*, *104*(12), 5241–5245. https://doi.org/10.1073/pnas.0608222104

Kirby, S., Griffiths, T., & Smith, K. (2014). Iterated learning and the evolution of language. *Current Opinion in Neurobiology*, *28*, 108–114. https://doi.org/10.1016/j.conb.2014.07.014

Kong, E. J., & Edwards, J. (2016). Individual differences in categorical perception of speech: Cue weighting and executive function. *Journal of Phonetics*, *59*, 40–57. https://doi.org/10.1016/j.wocn.2016.08.006

Kuznetsova, A., Brockhoff, P. B., & Christensen, R. H. B. (2017). lmerTest package: Tests in linear mixed effects models. *Journal of Statistical Software*, *82*(13), 1–26. https://doi.org/10.18637/jss.v082.i13

Labov, W. (1979). Locating the frontier between social and psychological factors in linguistic variation. *Individual Differences in Language Ability and Language Behavior*, 327–340. https://doi.org/10.1016/b978-0-12-255950-1.50025-1

Labov, W. (1986). The social origins of sound change. In H. B. Allen & M. D. Linn (Eds.), *Dialect and language variation* (pp. 524–541). Elsevier. https://doi.org/10.1016/B978-0-12-051130-3.50044-6

Labov, W. (1994). *Principles of linguistic change: Vol. 1: Internal* factors. Oxford: Blackwell.

Labov, W. (2001). *Principles of linguistic change: Vol. 2: Social* factors. Blackwell.

Lev-Ari, S., & Peperkamp, S. (2013). Low inhibitory skill leads to non-native perception and production in bilinguals' native language. *Journal of Phonetics*, *41*(5), 320–331. https://doi.org/10.1016/j.wocn.2013.06.002

Lev-Ari, S., & Peperkamp, S. (2014). The influence of inhibitory skill on phonological representations in production and perception☆. *Journal of Phonetics*, *47*, 36–46. https://doi.org/10.1016/j.wocn.2014.09.001

Lindsey, D. T., & Brown, A. M. (2002). Color naming and the phototoxic effects of sunlight on the eye. *Psychological Science*, *13*(6), 506–512.

Luengo Kanacri, B. P., Eisenberg, N., Tramontano, C., Zuffiano, A., Caprara, M. G., Regner, E., Zhu, L., Pastorelli, C., & Caprara, G. V. (2021). Measuring prosocial behaviors: Psychometric properties and cross-national validation of the prosociality scale in five countries. *Frontiers in Psychology*, *12*, 693174. Retrieved from https://www.frontiersin.org/articles/10.3389/fpsyg.2021.693174

Lupyan, G., & Dale, R. (2010). Language structure is partly determined by social structure. *PLoS ONE*, *5*(1), e8559. https://doi.org/10.1371/journal.pone.0008559

MacKeigan, T., & Muth, S. Q. (2006). A grammatical network of Tzotzil-Mayan colour terms. In C. P. Biggam & C. Kay (Eds.), *Progress in colour studies* (pp. 25–36). John Benjamins Publishing Company. https://doi.org/10.1075/z.pics1.06mac

Majid, A., & Kruspe, N. (2018). Hunter-gatherer olfaction is special. *Current Biology*, *28*(3), 409–413.e2. https://doi.org/10.1016/j.cub.2017.12.014

Mielke, J., Baker, A., & Archangeli, D. (2016). Individual-level contact limits phonological complexity: Evidence from bunched and retroflex /ɹ/. *Language*, *92*(1), 101–140. https://doi.org/10.1353/lan.2016.0019

Moisik, S. R., & Dediu, D. (2017). Anatomical biasing and clicks: Evidence from biomechanical modeling. *Journal of Language Evolution*, *2*(1), 37–51. https://doi.org/10.1093/jole/lzx004

Mothe, C., & Nguyen-Thi, T. U. (2021). Does age diversity boost technological innovation? Exploring the moderating role of HR practices. *European Management Journal*, *39*(6), 829–843. https://doi.org/10.1016/j.emj.2021.01.013

Mudd, K., de Vos, C., & de Boer, B. (2022). Shared context facilitates lexical variation in sign language emergence. *Languages*, *7*(1), 31. https://doi.org/10.3390/languages7010031

Natale, M. (1975). Convergence of mean vocal intensity in dyadic communication as a function of social desirability. *Journal of Personality and Social Psychology*, *32*(5), 790–804. https://doi.org/10.1037/0022-3514.32.5.790

Navarro, D. J., Perfors, A., Kary, A., Brown, S. D., & Donkin, C. (2018). When extremists win: Cultural transmission via iterated learning when populations are heterogeneous. *Cognitive Science*, *42*(7), 2108–2149. https://doi.org/10.1111/cogs.12667

Nettle, D. (2012). Social scale and structural complexity in human languages. *Philosophical Transactions of the Royal Society B: Biological Sciences*, *367*(1597), 1829–1836. https://doi.org/10.1098/rstb.2011.0216

Newman, R. (1997). Individual differences and the link between speech perception and speech production. *Journal of the Acoustical Society of America*, *102*, (5_Supplement), 3114. https://doi.org/10.1121/1.420560

Newman, R. S., Clouse, S. A., & Burnham, J. L. (2001). The perceptual consequences of within-talker variability in fricative production. *Journal of the Acoustical Society of America*, *109*(3), 1181–1196. https://doi.org/10.1121/1.1348009

Niederhoffer, K. G., & Pennebaker, J. W. (2002). Linguistic style matching in social interaction. *Journal of Language and Social Psychology*, *21*(4), 337–360. https://doi.org/10.1177/026192702237953

Nielsen, M. W., Bloch, C. W., & Schiebinger, L. (2018). Making gender diversity work for scientific discovery and innovation. *Nature Human Behaviour*, *2*(10), 726–734. https://doi.org/10.1038/s41562-018-0433-1

Noble, B., & Fernandez, R. (2015). Centre stage: How social network position shapes linguistic coordination. In *Proceedings of the 6th Workshop on Cognitive Modeling and Computational Linguistics* (pp. 29–38). https://doi.org/10.3115/v1/W15-1104

Nölle, J., Fusaroli, R., Mills, G. J., & Tylén, K. (2020). Correction: Language as shaped by the environment: Linguistic construal in a collaborative spatial task. *Palgrave Communications*, *6*(1), 44. https://doi.org/10.1057/s41599-020-0421-8

Pakendorf, B., Dobrushina, N., & Khanina, O. (2021). A typology of small-scale multilingualism. *International Journal of Bilingualism*, *25*(4), 835–859. https://doi.org/10.1177/13670069211023137

Pardo, J. S. (2006). On phonetic convergence during conversational interaction. *Journal of the Acoustical Society of America*, *119*(4), 2382–2393. https://doi.org/10.1121/1.2178720

Peterson, G. E., & Barney, H. L. (1952). Control methods used in a study of the vowels. *Journal of the Acoustical Society of America*, *24*(2), 175–184. https://doi.org/10.1121/1.1906875

Postmus, T. (2023). *A review on linguistic style matching*. University of Twente. Retrieved from https://essay.utwente.nl/94402/7/Postmus_BA_EEMCS.pdf

R Core Team. (2023). *R: A language and environment for statistical computing*. R Foundation for Statistical Computing. Retrieved from https://www.R-project.org/

Raviv, L., & Arnon, I. (2018). Systematicity, but not compositionality: Examining the emergence of linguistic structure in children and adults using iterated learning. *Cognition*, *181*, 160–173. https://doi.org/10.1016/j.cognition.2018.08.011

Raviv, L., Meyer, A., & Lev-Ari, S. (2019a). Compositional structure can emerge without generational transmission. *Cognition*, *182*, 151–164. https://doi.org/10.1016/j.cognition.2018.09.010

Raviv, L., Meyer, A., & Lev-Ari, S. (2019b). Larger communities create more systematic languages. *Proceedings of the Royal Society B: Biological Sciences*, *286*(1907), 20191262. https://doi.org/10.1098/rspb.2019.1262

Raviv, L., Meyer, A., & Lev-Ari, S. (2020). The role of social network structure in the emergence of linguistic structure. *Cognitive Science*, *44*(8), e12876. https://doi.org/10.1111/cogs.12876

Regier, T., Carstensen, A., & Kemp, C. (2016). Languages support efficient communication about the environment: Words for snow revisited. *PLOS ONE*, *11*(4), e0151138. https://doi.org/10.1371/journal.pone.0151138

Rogers, R. D., & Monsell, S. (1995). Costs of a predictible switch between simple cognitive tasks. *Journal of Experimental Psychology: General*, *124*(2), 207–231. https://doi.org/10.1037/0096-3445.124.2.207

Sandler, W., Meir, I., Padden, C., & Aronoff, M. (2005). The emergence of grammar: Systematic structure in a new language. *Proceedings of the National Academy of Sciences*, *102*(7), 2661–2665. https://doi.org/10.1073/pnas.0405448102

Scharf, S., Foster, J., & Behdinan, K. (2014). Optimizing linguistic diversity in highly multicultural engineering design teams. In *2014 ASEE Annual Conference & Exposition Proceedings* (pp. 24.959.1–24.959.11). https://doi.org/10.18260/1-2–22892

Schick, J., Fryns, C., Wegdell, F., Laporte, M., Zuberbühler, K., van Schaik, C. P., Townsend, S. W., & Stoll, S. (2022). The function and evolution of child-directed communication. *PLoS Biology*, *20*(5), e3001630. https://doi.org/10.1371/journal.pbio.3001630

Scott-Phillips, T. C., & Kirby, S. (2010). Language evolution in the laboratory. *Trends in Cognitive Sciences*, *14*(9), 411–417. https://doi.org/10.1016/j.tics.2010.06.006

Sebregts, K. (2014). *The sociophonetics and phonology of Dutch r*. LOT, Netherlands Graduate School of Linguistics.

Simmons-Mackie, N. (2018). Communication partner training in aphasia: Reflections on communication accommodation theory. *Aphasiology*, *32*(10), 1215–1224. https://doi.org/10.1080/02687038.2018.1428282

Tamariz, M., & Kirby, S. (2016). The cultural evolution of language. *Current Opinion in Psychology*, *8*, 37–43. https://doi.org/10.1016/j.copsyc.2015.09.003

Thompson, B., Raviv, L., & Kirby, S. (2019). Complexity can be maintained in small populations: A model of lexical variability in emerging sign. *Talk presented at the Interaction and the Evolution of Linguistic Complexity Workshop*. Edinburg, UK.

van Baaren, R. B., Holland, R. W., Steenaert, B., & van Knippenberg, A. (2003). Mimicry for money: Behavioral consequences of imitation. *Journal of Experimental Social Psychology*, *39*(4), 393–398. https://doi.org/10.1016/S0022-1031(03)00014-3

Walter, H. (2016). La norme linguistique dans le dictionnaire de l'académie française. *La linguistique*, *52*(1), 55–68. https://doi.org/10.3917/ling.521.0055

Willemyns, M., Gallois, C., Callan, V. J., & Pittam, J. (1997). Accent accommodation in the job interview: Impact of interviewer accent and gender. *Journal of Language and Social Psychology*, *16*(1), 3–22. https://doi.org/10.1177/0261927x970161001

Woensdregt, M. S., Spike, M., de Haan, R., Wareham, T., van Rooij, I., & Blokpoel, M. (2021). Why is scaling up models of language evolution hard? In *Proceedings of the Annual Meeting of the Cognitive Science Society*, *43*(43). Retrieved from https://escholarship.org/uc/item/021734q4

Yu, A. C. L., & Zellou, G. (2019). Individual differences in language processing: Phonology. *Annual Review of Linguistics*, *5*(1), 131–150. https://doi.org/10.1146/annurev-linguistics-011516-033815

Zellou, G. (2017). Individual differences in the production of nasal coarticulation and perceptual compensation. *Journal of Phonetics*, *61*, 13–29. https://doi.org/10.1016/j.wocn.2016.12.002

Supporting Information

Additional supporting information may be found online in the Supporting Information section at the end of the article.

Supporting Information